

A computational framework for understanding decision confidence

A Dissertation Presented

By

Joshua I. Sanders

to

The Watson School of Biological Sciences

at Cold Spring Harbor Laboratory

in Partial Fulfillment of the Requirements

for the degree of

Doctor of Philosophy

in

Biological Sciences

September 2013

Contents

Acknowledgements	iv
List of Abbreviations	v
List of Figures	vi
Foreword	ix
1 Introduction	1
1.1 General properties of a decision confidence measure	3
1.1.1 Confidence predicts accuracy	4
1.1.2 Confidence reflects <i>perceptual</i> discriminability	6
1.1.3 Confidence reflects deliberation	7
1.2 Explicit decision confidence measures available in humans	9
1.3 Implicit decision confidence measures in humans and animals	10
1.3.1 Uncertain option	10
1.3.2 Decline option	11
1.3.3 Post-decision wager	12
1.3.4 Post-decision temporal wager	13
1.4 Models of perceptual discrimination and reaction time	14
1.4.1 Accumulator models	14
1.4.2 Drift diffusion models	16
1.5 Models of perceptual discrimination and confidence	18
1.5.1 Accumulator model	19
1.5.2 Coupled accumulator model	20
1.5.3 Reaction time model	23
1.5.4 Variance model	24
1.5.5 Two stage dynamic signal detection (2DSD) model	25
1.6 Head-fixed behaviors available for studies of perceptual discrimination in rodent	26
1.6.1 Advantages and disadvantages of the mouse model for behavior	27
1.6.2 GO/NOGO and 2AFC task designs	27
1.7 Outline of chapters	28
2 Methodology	31
2.1 Development of decision confidence reporting tasks for humans and rats	31
2.1.1 The random click stimulus	31
2.1.2 Human perceptual confidence task design	33
2.1.3 Rodent perceptual confidence task design	43
2.1.4 A quiz to study the role of sensory pipeline noise in confidence	48
2.1.5 Design of a real-time click generator and response capture device	52

2.2	Data analysis.....	55
2.2.1	Statistics	56
2.2.2	Model fitting and measuring goodness of fit.....	58
2.2.3	Reverse correlation analysis.....	59
2.2.4	ROC analysis to determine how well evidence predicts behavior	60
3	Verbal confidence recapitulates the brain's statistical confidence in its assertions ..	61
3.1	Normative confidence predicts interrelationships of confidence, evidence and accuracy	61
3.2	Statistical confidence measures are characterized by normative model predictions	65
3.3	Human verbal confidence in percept classifications approximates statistical confidence	68
3.3.1	Discriminability measures with free-response Poisson click evidence	69
3.3.2	Estimating the time of motor response commitment and motor response latency	72
3.3.3	Click train evidence was the major determinant of subject choices.....	75
3.3.4	Subjects adopted stable evidence sampling strategies.....	80
3.3.5	Verbal confidence reports reproduce predicted normative model patterns .	82
3.3.6	Verbal confidence properties are stable within and across sessions	84
3.4	Human confidence in in non-sensory classifications approximates statistical confidence	86
3.4.1	Psychometric and chronometric performance on the Confography task.....	87
3.4.2	Verbal confidence in general knowledge reproduces normative model patterns	89
3.5	Conclusions.....	91
4	Post-decision time investment provides access to decision confidence in two species	93
4.1	Normative confidence patterns characterize the time investment measure in human	93
4.1.1	Theoretical relationship of time investment to abstract confidence	94
4.1.2	Human subjects perform the intended time investment task	95
4.1.3	Human time investment produces patterns expected of a confidence measure	97
4.2	Human time investment is strongly correlated with verbal confidence	101
4.3	Rat time investment produces patterns expected of a confidence measure.....	103
4.4	Conclusions	106
5	The Opportunistic Coupled Accumulator (OCA) model of confidence	108
5.1	Methods	109
5.2	Reaction time models do not account for patterns in confidence reports	110
5.3	Variance models fail because variance in evidence is unrelated to confidence.....	113
5.4	Accumulator models predict an incorrect pattern in use of recent evidence	117
5.5	Post-decision models overestimate how well post-decision evidence predicts confidence	121
5.6	Coupled accumulator models cannot fit both recent evidence and confidence patterns	125
5.7	Our OCA model simultaneously fits interrelationships of five key behavior measures	128
5.8	Conclusions.....	136

6 A device to capture multiple choice decision responses in head-fixed mice	139
6.1 Design of the choice ball device and experimental apparatus	139
6.2 Design of the task and training protocol.....	142
6.2.1 Subjects and surgery.....	142
6.2.2 Task flow of events	143
6.2.3 Training protocol.....	145
6.2.4 Readout of choice responses: mice respond with rapid (~200ms) paw motions	147
6.3 Mice perform the task and show evidence of speed/accuracy tradeoff.....	148
6.3.1 Mice learn the task within weeks	148
6.3.2 Mice show psychometric discrimination.....	149
6.3.3 Mice sample difficult stimuli for longer, and are more accurate for longer RT	149
6.4 Mice reverse their choices for a performance gain.....	150
6.4.1 Readout of choice reversals.....	151
6.4.2 Mice reverse their choices to optimize performance.....	151
6.5 Conclusions	152
7 Conclusions and perspectives	155
Appendix I: Derivation of confidence signal properties from normative confidence ...	163
Appendix II: Bpod: a microcontroller based real-time behavior acquisition and control system	169
Appendix III: Decision confidence with an alternative evidence processing strategy: a case study.....	175
Appendix IV: Towards an identity test for confidence signals.....	179
References	183

Acknowledgements

For consistently sharing his insight and providing crucial guidance, I would foremost like to thank my research advisor, Adam Kepecs. The path to discovery is rarely as straightforward as it is cast in hindsight, and for this reason I am particularly indebted to Adam for his patience - during my mid-graduate school foray into tangentially relevant electrical engineering, and for persistence in calibrating my temporal discounting function to the pace of high quality scientific results and authorship. I am eternally grateful to my colleagues and de facto family, Duda Kvitsiani and Ebru Demir, for providing a solid anchor of social support when it was needed most, and for keeping the days interesting and unpredictable. I was extremely fortunate to receive academic mentorship from Bruce Stillman, who despite his obligations as CSHL President, met with me whenever I sought his counsel and ensured that my research was on track. I was also fortunate to work in a lab with a sense of common purpose and good will – for that, I thank my labmates in Kepecs Lab: Sachin Ranade, Hyun Pi, Junya Hirokawa, Alexander Vaughan and especially Balazs Hangya for his contribution of key derivations to the present work. I'd like to thank my building-mates Francesca Anselmi for being the kind of rare friend who's genuinely there when you stumble, and Florin Albeanu and Priyanka Gupta for keeping me in line. Cold Spring Harbor systems neuroscience labs share a uniquely open academic spirit. I am indebted to Anne Churchland for guidance and technical assistance in developing human behavior protocols, to Alexei Koulakov for contributing insight and a key formula that explains time investment to this thesis, to Gonzalo Otazu and Gidon Felsen for teaching me practical electrophysiology, to Masayoshi Murakami and Hatim Zariwala for crucial tips on scientific programming, and to Zach Mainen for his expert mentorship and for encouraging my independent research while I was his technician. For veteran technical mentorship and general wisdom, I am especially thankful to our senior research associate, Barry Burbach. I am grateful to my thesis committee members – Glenn Turner, Bruce Stillman, Anthony Zador, Carlos Brody and our external member Nathaniel Daw, for their insightful guidance as my research matured, and for helping me to maintain the crucial focus that brought my projects to completion. I sincerely thank the Watson School staff, Alyson Kass-Eisler, Dawn Meehan and Kim Creteur, under the leadership of Dean Leemor Joshua Tor and Dean Alexander Gann, for providing a well-structured and supportive learning environment. To the Gerry Family – I deeply appreciate your provision of research funding for my term as a graduate student, in an economy where such support is a genuine luxury. Finally, I am grateful to those in the greater Cold Spring Harbor and Stony Brook communities, for making my years of formal education a fruitful and enjoyable phase of life.

List of Abbreviations

2AFC	Two alternative forced choice
OCA	Opportunistic coupled accumulator
2DSD	Two-stage dynamic signal detection
ANOVA	Analysis of variance
AUC	Area under (ROC) curve
CAD	Computer aided design
DDM	Drift diffusion model
CI	Confidence interval
ITI	Inter-trial interval
LED	Light emitting diode
MOSFET	Metal oxide field effect transistor
OFC	Orbitofrontal cortex
PC	Personal Computer
PX	Pixels
QMP	Quantile maximum probability
RAM	Random access memory
ROC	Receiver operating characteristic
RT	Reaction time
SD	Standard deviation
SEM	Standard error of the mean
TI	Time investment
TTL	Transistor-transistor logic
USB	Universal serial bus

List of Figures

Chapter 1

1.1 Calibration scores for experimental conditions across studies	6
1.2 Measures of confidence in sequential sampling models of choice	19
1.3 Processing of evidence in coupled accumulator models	21

Chapter 2

2.1 Time course of the random click stimulus.....	33
2.2 Layout of the confidence response keypad.....	35
2.3 Flow of trial events for explicit-only and implicit tasks.....	37
2.4 Game elements to encourage alertness and long-term participation	41
2.5 Spending game currency does not affect block mean performance or confidence .	42
2.6 Rodent interface port configuration.....	44
2.7 Flow of trial events in the rat implicit confidence task	45
2.8 Confography subject instruction screen	50
2.9 Flow of trial events for Confography task.....	52
2.10 The Pulse Pal device and custom interface for human responses	54

Chapter 3

3.1 Three patterns in decision confidence derived from the normative model.....	63
3.2 Statistical confidence measures reproduce patterns predicted by the normative model	67
3.3 In a free response task, underlying rate does not capture the strength of Poisson evidence	70
3.4 The balance of experienced evidence is a continuum despite intended difficulty Categories	71
3.5 A measure of post-hoc discriminability, β captures the full range of graded Performance	72
3.6 Evidence in the final moments of sampling does not predict choice	73
3.7 Subjects responding as soon as 125ms after stimulus onset discriminate above Chance	75
3.8 Discrimination accuracy is stable and varies as a function of our discriminability Measure.....	76
3.9 Click stream evidence dominated choice, despite small side and reward-history Biases.....	79
3.10 Subjects sampled evidence strategically to economize time	81
3.11 Confidence reports predict outcome and contain insight into perceptual evidence	83
3.12 Subjects showed unique but consistent confidence levels.....	85
3.13 Population log ratio and RT measures support use of general knowledge to decide	88
3.14 Normative model predictions describe confidence in general knowledge	90
3.15 Temporal properties of confidence in general knowledge	91

Chapter 4

4.1	Subjects sampled Poisson click evidence to maximize performance	96
4.2	Human subjects invested liberally and invested more time when they were correct	98
4.3	Human time investments resemble normative model predictions for confidence...	99
4.4	Human time investments for individual subjects resemble normative patterns	100
4.5	Time investments are highly correlated with subsequent verbal confidence reports	102
4.6	Rats perform the time investment task and invest more conservatively than humans	105
4.7	Rat time investments resemble normative model predictions for confidence	106

Chapter 5

5.1	The reaction time model of confidence makes three incorrect predictions	112
5.2	Variability of Poisson evidence does not account for decision confidence	116
5.3	The accumulator model incorrectly predicts the reverse correlation of evidence ...	119
5.4	The human stopping rule filtered for supporting evidence and against contradictory evidence in the choice-informative final second of sampling	120
5.5	Confidence is computed using evidence gathered both during and after choice	124
5.6	The coupled accumulator model cannot simultaneously fit patterns in evidence history and confidence	126
5.7	No previous model correctly predicts temporal patterns of evidence contribution to choice and confidence	127
5.8	Illustration of OCA model variables and parameters in an example trial	130
5.9	The OCA model simultaneously fits patterns in human behavior that eluded previous models	131
5.10	For individual subject H4, the OCA model simultaneously approximates 12 key interrelationships between decision measures	133
5.11	OCA model fit for subject H3	134
5.12	OCA model fit for subject H2	134
5.13	OCA model fit for subject H1	136

Chapter 6

6.1	The choice ball apparatus	141
6.2	A two-alternative forced choice task for head-fixed mice	144
6.3	Rapid readout of choice responses	147
6.4	Mice learn the trackball task consistently and with high accuracy	148
6.5	Choice accuracy varies with discriminability	149
6.6	Reaction times correlate with stimulus difficulty and performance	150
6.7	Detection of choice reversals in mice	151
6.8	Choice reversals are used to correct errors	152

Appendix II

A2.1	Components of the Bpod interface circuit board	172
A2.2	Pseudo-code of Bpod firmware	173

Appendix III

A3.1 With strong evidence and biased, rapid judgments similar to our rats, normative patterns persist in explicit confidence.....	177
A3.2 The processes of gathering evidence for choice and for confidence are dissociable	178

Appendix IV

A4.1 Normative predictions are not reproduced by two global gain processes	181
--	-----

Foreword

We are on the cusp of a very exciting era in brain research; emerging molecular and brain-computer interface technologies increasingly enable us to directly observe and manipulate the physical mechanisms underlying mental processes. However, many of the brain's most interesting and elusive abilities depend on inferred mental variables which are difficult to capture experimentally. For instance, while *confidence* enables a brain to make shrewd investments, learn optimal strategies and succeed in social exchanges, it strongly manifests to conscious human observers as a *feeling* which is reported verbally. How does one identify something as amorphous as a feeling amongst the brain's electrical signals? What does a confidence signal look like? In this dissertation, I answer this fundamental question by relating human self-reports of the feeling of confidence to *statistical* confidence with a normative model that makes three strong predictions about how statistical confidence signals should look in different projections of data. I show that these signature patterns, one of which had not previously been reported in humans, robustly describe confidence in three markedly different contexts: verbal confidence reports from humans performing a range of decision making tasks, confidence-guided investment decisions in humans and rodents, and p-value confidence measures produced by common statistical hypothesis tests. With this understanding of what is being computed – an approximation of statistical confidence - I sought to determine which *algorithm* the brain uses to generate confidence reports. I developed a novel decision confidence reporting task, where confidence reports are based on high resolution, temporally structured evidence. Using these data, I was able to rule out five previously suggested algorithms, each predicting qualitatively different patterns in confidence and the history of evidence use. To account for our data, I developed a new algorithm – the Opportunistic Coupled Accumulator (OCA) - for computing choice and confidence from sensory evidence in real-time. I found that OCA is able to simultaneously fit patterns in human confidence reports and evidence use to within statistical error of our data in several key projections. The algorithm provides insight that may help relate measurements of functioning neural circuits to the emergence and usage of an abstract mental confidence variable. To enable observation of abstract mental variables like confidence in the brain, the brain must be engaged in the kind of sophisticated decision making behaviors that require abstract economic judgments. I developed a new device for capturing two-choice decisions in the head-fixed mouse, availing more sophisticated decision making behaviors to analysis by functional microscopy and related techniques. Our findings place mental confidence on the rigorous computational footing which is a necessary bridge from abstract variables to electrical impulses, and provide tools to study how confidence is physically computed in the neural circuits of humans and rodents.

Chapter 1

Introduction

Decision confidence is an estimate of the probability that a decision maker is correct, given the evidence used to decide. It is a key mental variable, necessary for the brain to optimally exploit imperfect information about the state of its environment (Knill and Pouget 2004). Despite its overt similarity to the notion of confidence in statistical hypothesis testing, precisely indicating decision confidence remains a semantic challenge in neuroscience and cognitive science literature. Decision confidence is often introduced as a feeling (Insabato, Pannunzi et al. 2010), through examples (Moreno-Bote 2010) or else entirely recast as “certainty” (Fimbel, Michaud et al. 2009), “subjective probability” (Kahneman and Tversky 1972) or “introspective accuracy” (Fleming, Weil et al. 2010). The lack of a unified language to indicate such a fundamental operation of the mind despite over a century of research illustrates the degree to which a clarifying framework is necessary, built upon a precise quantitative definition of decision confidence.

The literature on modeling decision confidence in perceptual discrimination is surprisingly sparse – as recently as 2001, theorist Douglass Vickers noted that in the cognitive psychology of decision making, confidence is “relied on for its usefulness, but overlooked as an interesting variable in its own right” (Vickers 2001). However, research seeking to explain how organisms generate confidence estimates has experienced a resurgence in the past decade e.g. (Kepecs and Mainen 2012) – perhaps fueled by the growing appreciation that confidence estimates of various types confer an advantage to machine learning algorithms faced with similar perceptual challenges (Schapire and Singer 1999; Schwenk and Bengio 2000; Sollich 2002).

A precise understanding of how the brain computes confidence is central to an understanding of how the brain functions. When equated terminologically with the Bayesian posterior of the decision process (an estimate of the probability that an option is correct given evidence), confidence is already acknowledged in neuroscience literature as a key ingredient in estimating the value of

outcomes (Behrens, Woolrich et al. 2007), learning (Courville, Daw et al. 2006) and perceptual decision making (Beck, Ma et al. 2008; Rao 2010). Knowledge of the brain's mechanism for computing decision confidence will lend insight into how the brain accomplishes these important functions, and what goes wrong in diseases where they are impaired.

Our research was inspired by previous findings by Kepecs et al, that the firing patterns of neurons in the orbitofrontal cortex of rat show characteristics of a confidence signal (Kepecs, Uchida et al. 2008). In the same study, confidence patterns emerged in behavior when rats were permitted to save time by strategically re-initiating trials. In order to describe how confidence signals captured from the brain are used to drive behavior, we needed a quantitative model describing how a measure of behavior corresponds to an animal's internal representation of decision confidence. More primarily, confidence is an abstract mental function traditionally associated with humans and to a lesser extent, lower primates. The burden of proof was upon us to establish that a rodent can be confident at all. While our first experiments showed that in a human analog of our rodent discrimination task, temporal wagers are strongly correlated with explicit confidence reports, we soon realized that even human explicit reports have not been captured by a model of discrimination and confidence that reproduces all of the patterns we observed in our data.

This chapter begins by establishing the historical precedent for our profile of a confidence signal - three patterns that all representations of abstract confidence should match. I review ways in which an experimenter can measure an organism's confidence, and how these measures might correspond to the abstract mental confidence value they are derived from. I then summarize prior algorithms describing how decision makers determine their choices and confidence. Looking beyond the algorithm, we would like to determine the *mechanism* used by the brain to compute confidence at the level of neural circuits. I review the challenges of studying the circuits underlying decision making with powerful new optogenetic and neuroimaging tools that require a head-fixed behavioral assay.

1.1 General properties of a decision confidence measure

To identify a confidence signal in the brain, or to be aware when a subject who claims to be reporting confidence is actually reporting something else, we first need to know what a confidence report looks like. The confidence reports produced by a human decision maker have characteristic relationships to other common measures of decision making – accuracy, the quality of evidence used and evidence sampling time / reaction time. However, which patterns to expect from *all* confidence measures has not been established a priori. In practice, researchers rely on collections of prior generalizations about the empirical properties of confidence reports (Vickers 2001). To establish the fitness of a confidence measure beyond noting that its patterns resemble other measures, a clear set of predictions derived from a normative definition of decision confidence can establish which patterns to expect a priori. We provide this derivation in Appendix I, and argue for a profile of three patterns: that confidence predicts accuracy, that confidence reflects the strength of perceptual evidence, and that confidence informs accuracy despite a fixed level of external evidence. The first pattern has been consistently observed in confidence literature. The second pattern has been observed for pooled choice outcomes, and a slight positive correlation between confidence and evidence strength was sparsely reported. However, when separated by correct discrimination and errors, we found a pattern in errors that manifests in a wide array of confidence reports and had been ignored in literature prior to Kepecs et al 2008 – the lowest possible confidence occurs in the presence of the strongest absolute evidence. The third pattern, that confidence predicts accuracy for fixed evidence strength, had also been ignored. In the following three sections, I review a century of prior research characterizing the relationship of confidence to other decision measures.

1.1.1 Confidence predicts accuracy

The fact that human confidence reports can forecast the accuracy of choices between two alternatives was first formalized over a century ago (Jastrow and Peirce 1884). Interchangeably termed “calibration” or “realism” (Lichtenstein, Fischhoff et al. 1981), the degree to which confidence reports predict outcome probabilities has been used as the field’s de facto measure of report fitness. Since calibration informs human calculation of economic risk (Kahneman and Tversky 1979), finding patterned ways in which humans forecast inaccurately has been the focus of a considerable amount of research in economics.

Calibration in a decision making task is a measure derived from the Brier score (Lichtenstein and Fischhoff 1980; Keren 1991; Baranski and Petrusic 1994; Koehler and Harvey 1997; Kvidera and Koutstaal 2008), a proper scoring rule for binary events originally introduced to compare the effectiveness of forecasting measures in meteorology (Brier 1950). While generally applicable to predictive measures, the Brier score for decision confidence reporting is given as:

$$B = \frac{1}{n} \sum_{t=1}^n (c - o)^2 \quad (1.1)$$

where c is the subject’s confidence formulated as the predicted probability of a positive outcome, and o is the outcome (0 or 1), on trial t . Most studies of decision confidence use a confidence reporting scale with discrete report categories. A calibration score, C_s , is then evaluated by the following related formula (Lichtenstein and Fischhoff 1977; Björkman, Juslin et al. 1993):

$$C_s = \frac{1}{n} \sum_{t=1}^T n_t (c_t - o_t)^2 \quad (1.2)$$

where n is the total number of trials, T is the number of divisions of the reporting scale, n_t is the number of times the t^{th} division of the scale was selected in the session, c_t is the expected mean outcome probability for perfect calibration (derived from the definition of the boundaries of category

t) and o_t is the actual mean outcome for all trials in category t . Calibration is thus a weighted mean of the difference between predictions and observations for each category on a confidence reporting scale. A calibration score of 0 implies that confidence in each category perfectly predicts outcome probability, while a score of 1 indicates that confidence reports have no predictive power. The score does not indicate cases where *part* of the scale is used sub-optimally, and cannot distinguish between overconfidence and under-confidence.

To measure overconfidence or under-confidence, equation 1.2 is frequently used, though without squaring the term $c_{im} - o_t$ (Keren 1991; Jonsson and Allwood 2003). When subjects choose based on outcome history and/or general knowledge, the most consistently shown inaccuracies are overconfidence on difficult judgments and under-confidence on easy judgments (for a review of biases in non-perceptual confidence reports, see Lichtenstien, Fischhoff et al. 1981). However, when subjects choose based on classification of perceptual evidence, experimental findings are considerably more conflicted. Some perceptual discrimination studies find systematic under-confidence (Keren 1988; Björkman, Juslin et al. 1993; Olsson and Winman 1996; Stankov 1998), others find over-confidence (Kvidera and Koutstaal 2008) and still others find that subjects can be overconfident or under-confident depending on difficulty as for non-perceptual judgments (Griffin and Tversky 1992; Baranski and Petrusic 1994; Bar-Tal, Sarid et al. 2001).

Despite the range of methods in these perceptual tasks and disagreement about the direction of bias, the amount by which confidence reports diverge from perfect calibration in perceptual discrimination studies is consistently small. Below, I collected the range of typical calibration values for studies that provide calibration scores of separate experimental conditions and do not give an explicit speed instruction to their subjects (Figure 1.1).

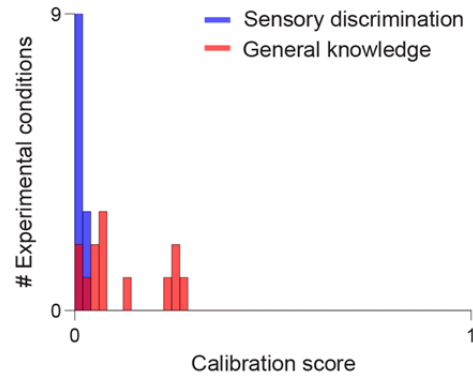


Figure 1.1: Calibration scores for experimental conditions across studies. Each data point in the histogram represents the calibration score for an experimental condition reported. Sensory discrimination studies (drawn in blue) used experimenter-provided evidence as the basis for choice and confidence. In general knowledge studies (drawn in red), subjects were quizzed on general knowledge topics (vocabulary, the population of cities, etc.) and provided confidence reports for their responses. Only studies of healthy subjects that provided calibration scores computed with equation 1.2 were included. Sensory discrimination studies are: (Keren 1988; Björkman, Juslin et al. 1993; Baranski and Petrusic 1994; Winman, Juslin et al. 1998; Merkle and Van Zandt 2006). General knowledge studies are: (Lichtenstein and Fischhoff 1980; Koehler and Harvey 1997; Bornstein and Zickafoose 1999; Jonsson and Allwood 2003).

This overview of calibration scores is not exhaustive, and is only intended to provide an idea for the range of calibration values typically reported. These summary findings indicate that in reporting confidence, human decision makers have conscious access to an intriguing mental variable that *explicitly predicts the likelihood of future events* (outcomes). Especially in judgments where sensory evidence directly informs discrimination, confidence can be a nearly perfect likelihood estimate (Figure 1.1). How does the brain make these predictions? Our normative model in section 3.1.1 suggests that confidence is computed using a mental estimate of the quality of the evidence used by the brain for choice. If correct, prior studies with graded discriminability should show that in addition to choice accuracy, confidence reflects perceptual evidence.

1.1.2 Confidence reflects *perceptual* discriminability

While the relation of confidence calibration to evidence strength has been studied extensively, how confidence *reports* directly correspond to evidence strength has received considerably less attention. A positive relationship between confidence reports and experienced

evidence was originally suggested by Charles S. Peirce in a philosophical investigation of subjective probabilities (Peirce 1877). A subsequent investigation showed that confidence reports entered on a 5-point scale, correlate with discriminability in a line length discrimination task (Vickers and Packer 1982). The study compared conditions with speed and accuracy instructions, showing that the correlation between confidence and evidence was stronger in the accuracy condition. Another study also using line length discrimination showed a weaker effect (Baranski and Petrusic 1998). Accuracy on the easiest and most difficult conditions tested ranged from 75% to 88%, corresponding to a small but significant difference of 0.2 in confidence on the 4-point scale.

While confidence indeed correlates with the experimentally intended strength of evidence on correct trials in our data, it is anti-correlated with evidence on error trials such that *the lowest confidence reports are generated in the presence of the evidence intended to be strongest*. This counterintuitive trend in errors was first reported by Kepecs et al in the cells of rat orbitofrontal cortex, and in the probability that a rat will “reinitiate” a trial while waiting for a delayed reward (Kepecs, Uchida et al. 2008). The present research provides the first investigation of these trends in human explicit confidence reports (Section 3.1.4), and in several models of decision confidence (Chapter 5). This counterintuitive pattern has also been predicted by an artificial two-layer network of spiking neurons performing perceptual classification (Insabato, Pannunzi et al. 2010), and is a property of p-values in common two-sample statistical tests (section 3.2).

1.1.3 Confidence reflects deliberation

In our definition of confidence as a probability estimate conditioned on evidence, evidence is time invariant. Real-world decision makers sample evidence in time, and are often free to choose when to stop sampling (see section 1.4). Early researchers established that reaction time varies inversely with experimentally controlled strength of evidence in a sensory detection task (Cattell 1886). That an inverse correlation between confidence and reaction time exists has also been long appreciated (Henmon 1911; Volkmann 1934), consistent with the idea that confidence is derived from

the strength of perceptual evidence. Anti-correlation with reaction time has even been suggested as a general property of confidence reports (Vickers 2001). However, several studies have reported circumstances under which confidence and reaction time are dissociable.

Using a line length discrimination task, Baranski and Petrusic (1998) found that giving subjects a speed instruction virtually eliminated the correlation between confidence and reaction time, while this correlation was preserved for subjects with an accuracy instruction.

Merkle and Van Zandt (2006), found that reaction time and confidence were correlated for half of the reporting scale used. Using a visual discrimination task where subjects were required to determine whether more or fewer dots were shown on a screen with respect to an experimentally imposed category boundary, confidence was only correlated with reaction time when confidence was between its lowest value (0.5) and the middle of its range (0.75) on their 10-division reporting scale. Reaction times for confidence reports between 0.75 and 1 were indistinguishable.

Ascher (1979) found that confidence was *positively* correlated with reaction time, in a line length discrimination task where choices were entered on a 6 point bipolar scale, indicating high, medium or low confidence for each choice (1-3 = Strong...weak choice of hypothesis A, 4-6 = weak...strong choice of hypothesis B). This result raises the possibility that reaction time may be partially determined by the process of generating a confidence report. In a subsequent visual shape size discrimination task, subjects required to report confidence after choice had longer reaction times than subjects who were only required to discriminate. (Petrusic and Baranski 2003). The difference between the confidence and no-confidence reaction times was largest for easy trials, and smallest for difficult trials, suggesting that preparing for a *post-decisional* confidence report during choice is an active process that depends on the strength of evidence and contributes to reaction time.

In section 1.5.2, we will review a model of confidence and choice where confidence is determined causally by a quantitative assessment of reaction time. The fact that confidence has been shown to be dissociable from reaction time under some circumstances, casts serious doubt on the generality of this class of models. Additionally, these dissociations challenge the previously noted

presumption by Vickers et al (2001) that the confidence/reaction time anti-correlation is a general property of confidence reports.

1.2 Explicit decision confidence measures available in humans

The subjective experience of confidence contains a sense of magnitude, yet in providing a verbal confidence report, a subject must map this internal sense to some reporting scale designated by the experimenter. It is most straightforward to determine the calibration of confidence when subjects provide an explicit estimate of the probability that they are correct, often in the form of a percentage estimate (Lichtenstein, Fischhoff et al. 1981). With 51 divisions at single percentage resolution, the percent scale is closer to a true continuous scale than the 5-point scale we chose for our research. While studies directly comparing discretized and continuous confidence scales are scarce, one study of radiologists' confidence in their diagnoses showed that use of a discretized 5-point scale does not distort calibration (Rockette, Gur et al. 1992). More generally, the question of how response scale resolution affects measurement has been most thoroughly studied for the bipolar Likert scale (Likert 1932), a discrete scale used to capture self-reports of graded agreement or disagreement with a statement. A recent study of market surveys addressing the comparability of mean, variance, skewness and kurtosis for responses on Likert scales containing 5, 7 or 10 response divisions found their mean and variance statistics comparable with the only significant difference being a slightly lower mean on the ten point scale (Dawes 2008). Since decision confidence scales are usually unipolar in two-choice discrimination tasks (confidence in the chosen hypothesis), the 5 point confidence scale provides the same resolution in our range of interest as a bipolar Likert scale with 9 options (4 graded positive choices and one neutral).

The vast majority of studies specify a discretized confidence scale of 4-7 divisions, without disclosing the logic behind their choice. However, scales with more divisions are sometimes used. Adams and Adams (1961) use a verbally reported single percentage scale, with two justifications – that “performance can be compared with expectation at all points on the scale” (ignoring fractional

percentages) and that “individual differences on the meanings of different points on the scale are reduced” (Adams and Adams 1961) p.37. The authors do not cite evidence that individual differences in scale interpretation with fewer divisions can affect calibration. Despite the loss of resolution, reporting scales with as few as two divisions are common. Rahnev et al. (2012), in a study on the effect of visual cortical TMS on confidence reports, uses a two division scale “to keep the task as simple as possible and because a continuous confidence scale would have complicated our signal detection theoretic modeling” (Rahnev, Maniscalco et al. 2012) p1557.

1.3 Implicit decision confidence measures in humans and animals

While humans have the unique ability to consciously experience and verbally report their confidence, this type of report is not typical of how confidence is used by organisms to guide behavior. Confidence is used to inform risk taking, and in this role it is easier to isolate experimentally than in its other roles regulating learning or emotion. A variety of behavioral tasks with risk taking decisions have been developed to measure confidence. Since a verbal report is not requisite, many of these tasks can be used to measure confidence in animal models as well as humans.

1.3.1 Uncertain option

Early studies of implicit confidence employed two-choice discrimination tasks with the added option to indicate uncertainty (for a small but guaranteed reward) instead of classifying the stimulus. Preferential selection of the uncertain option on difficult discriminations was considered a signature of confidence. While originally employed in human studies (Watson, Kellogg et al. 1973), studies using uncertain option tasks have found use of confidence in monkeys (Shields, Smith et al. 1997; Smith, Shields et al. 1997), dolphins (Smith, Schull et al. 1995) and pigeons (Sole, Shettleworth et al. 2003). Uncertain option tasks have an important caveat; since the subjects receive more reward on average for choosing the uncertain option than for classifying when presented with difficult stimuli, the task can be solved by pairing the more difficult stimuli with the uncertain response using

reinforcement learning. In this scheme, a separate confidence estimate is not necessary and the task is in essence, a three choice discrimination task (Kepecs and Mainen 2012). To address this caveat, Smith et al designed a variant of the task where rewards were given following blocks of trials, based on performance within each block (Smith, Beran et al. 2006). However, the proper associations could still be learned by reinforcement learning paradigms with more sophisticated credit assignment procedures (Sutton and Barto 1998). More recent studies have generally employed confidence tasks with more sophisticated ways of isolating confidence.

1.3.2 Decline option

Decline option tasks are a variant of uncertain option tasks that provide a stimulus to classify and then two choices in series: 1. whether to solve the discrimination or decline for a small, certain reward, and 2. discrimination (if selected instead of the decline option). To control for whether the decline choice is simply paired with difficult stimuli, “forced” trials skip choice 1, and force the subject to enter their discrimination. If the decline choice was simply paired with difficult stimuli, for a given stimulus difficulty, accuracy on forced trials would not be different from accuracy on trials with the decline option. Rather, diminished accuracy on forced trials has been shown in monkeys (Hampton 2001; Kiani and Shadlen 2009) and rats (Foote and Crystal 2007), but not pigeons (Inman and Shettleworth 1999; Sutton and Shettleworth 2008), leaving open the possibility that acting based on confidence is an ability unique to mammals.

Decline option tasks depend upon differential accuracy between forced trials and trials with the option to decline. However, the same differential accuracy would be expected if subjects preferentially declined when they were aware of their own inattention – a process distinct from confidence (Kepecs and Mainen 2012). Moreover, the binary resolution of confidence (Solve=high, Decline=low) is not ideal for more sophisticated analyses of confidence (for instance, calibration or chronometry). A further shortcoming of both decline option tasks and uncertain option tasks is that

individual “uncertain” or “decline” trials do not also provide choices, prohibiting trial by trial analysis of the relationship between choice evidence and decision confidence.

1.3.3 Post-decision wager

One way to acquire both a choice and a confidence judgment on a single trial is to present these choices in series, as is generally done in human verbal decision confidence studies. An implicit equivalent of this is the post decision wager class of tasks. On each trial, the subject discriminates between two options, and subsequently casts a wager on their presumed reward. A post-decision wager task was used in humans by Persaud et al to study use of confidence information in primary cortical blindsight patient GY (Persaud, McLeod et al. 2007). On each trial, the subject was asked to indicate whether a grating pattern had been presented, and to wager either 50p or £1 (low or high) on the hypothesis that he had indicated correctly. Wagers on stimuli presented in GY’s normal visual field were virtually all high for correct trials (24 trials high, 1 trial low, 96% detection accuracy) while wagers for correct trials in his blind hemifield were cast randomly (67 high, 74 low) despite 70% accuracy, indicating that blindsight wagers were not calibrated to accuracy. While these results are consistent with a role for visual cortical processing in calculating confidence, the report is also a cautionary example of how design of a payoff matrix in a wagering task can be critical for interpreting wagers as confidence judgments. As the task was designed, if accuracy in this task is above chance, the optimal strategy is to always wager high (Clifford, Arabzadeh et al. 2008), making it difficult to discern between a suboptimal wagering policy and a corrupted confidence signal (Kepecs and Mainen 2012). Subsequently, Fleming et al. established that human subjects casting post-decision wagers in this task display loss aversion (Fleming and Dolan 2010). Loss aversion is the well-established tendency of humans to irrationally avoid risk in economic decisions (Kahneman and Tversky 1979), a factor that further complicates the isolation of a confidence signal from the discrete wager measure. A more sophisticated experimental design utilizing a continuous wagering scale tailored to each subject’s performance could potentially resolve these problems (Kepecs and Mainen

2012), providing a useful implicit confidence measure. Middlebrooks and Somer et al reported that rhesus monkeys can learn to bet strategically in a post-decision wager task using a discrete wagering scale of “high” and “low” options affecting reward probability (Middlebrooks and Sommer 2011). However, a task requiring a discrete post-decision wager of individual reward amount has yet to be established in rodents.

1.3.4 Post-decision temporal wager

In the context of a free response behavioral task organized as a session of individual trials, subjects seek to maximize their reward *rate* (Simen, Contreras et al. 2009). In a wagering task, a subject can affect reward rate (reward amount / time) by gambling to either maximize reward amount or minimize time invested per reward. A simple manipulation of two-choice discrimination that intuitively encourages a subject to gamble with their time is to delay reward delivery by a random interval once the subject has entered their response. While waiting for reward, the subject has the option to stop waiting and initiate the next trial at any time, making the optimal time investment proportional to their confidence in their choice. This style of confidence report was first used by Kepecs et al. to measure confidence in rats (Kepecs, Uchida et al. 2008), and is the implicit confidence measure characterized further in the present study. We show in chapter 4 that the post-decision temporal wager (termed the “time investment” (TI) measure in all further references) can be used to compare confidence between human and rodent at the resolution of individual trials. However, to extract an abstract confidence value from a time investment measure, three additional transformations must be accounted for beyond computing a simple post-decision wager – the subject’s knowledge of the reward delay distribution, the subject’s cost-function for time and the subject’s imprecision in estimation of time. The relationship between abstract confidence and time investment is explored in the beginning of Chapter 4, followed by a demonstration that time investment can be lawfully derived from a confidence signal and matches patterns in human self-reports of confidence. The availability of a task providing both a choice and a confidence measure on

the same trial in rodent models, provides researchers with access to this important decision making variable in a context amenable to recording and manipulation of neural circuitry.

1.4 Models of perceptual discrimination and reaction time

In order to understand how the brain computes decision confidence at the level of neural circuits, it is instructive to first determine which algorithm it uses to assemble evidence for separate options and choose among them. Noisy accumulation of evidence for each option is typically modeled as a sequential sampling process (Wald 1947). Evidence is accumulated by a separate decision variable for each choice, until a stopping rule terminates the sampling process. Models of choice implement this process either using discrete time steps (LaBerge 1962), or in continuous time (Audley and Pike 1965). Different stopping rules characterize the two major classes of sequential sampling model: “accumulator” (or in more recent literature, “race”) models terminate accumulation when the *absolute* evidence collected for an option exceeds a fixed threshold, and “diffusion” models terminate when the *difference* in evidence accumulated among options exceeds threshold (Ratcliff and Smith 2004). To differing degrees, specific implementations of these models can explain psychometric performance, reaction time distributions and speed/accuracy tradeoff. With additional provisions reviewed in section 1.5, both classes of model can generate confidence reports. More recent models have proposed more sophisticated stopping rules that have mixed absolute and relative characteristics (Moreno-Bote 2010) or take the subjective cost of time into account (Drugowitsch, Moreno-Bote et al. 2012).

1.4.1 Accumulator models

The earliest accumulator-class model of choice and reaction time was proposed by LaBerge et al as part of his recruitment theory of behavior (LaBerge 1962). This model operates by separately computing noisy sums of evidence supporting each choice at discrete time intervals, and terminates the choice process when the amount of evidence collected for one hypothesis exceeds a fixed

threshold. Noise sources affect the count of each counter as they progress towards threshold, and are independent for each count. Accumulation of evidence would proceed as follows for decision variables L and R:

$$\begin{aligned} L_{i+1} &= L_i + E_L + \eta_{L,i}; \\ R_{i+1} &= R_i + E_R + \eta_{R,i}; \\ \max(L_i, R_i) &< T \end{aligned} \tag{1.3}$$

where E_L and E_R are the mean strengths of evidence for left and right hypotheses respectively, η_L and η_R are noise vectors drawn independently from a standard Gaussian distribution with a free variance parameter, and T is a second free parameter; the choice threshold. Vickers et al improved upon this model, showing that a better reaction time fit could be achieved by drawing momentary evidence for each counter from separate Gaussian distributions for each hypothesis, where mean is equal to hypothesis evidence strength (Vickers 1979; Smith and Vickers 1988). An alternative variant, the Poisson race model, accrues discrete evidence for each hypothesis at random exponential intervals on a continuous time scale, for an improvement of reaction time fit under some conditions (Audley and Pike 1965; Van Zandt, Colonius et al. 2000).

Accumulator models have been used to model brain function during behavior. Most notably, an accumulator model applied to random dot motion discrimination in monkeys was used to explain firing of choice-selective neurons in area LIP, which ramp to an apparent fixed threshold at the moment of choice (Churchland, Kiani et al. 2008). However, drift diffusion models or “coupled accumulator” accumulator models with lateral inhibition (McClelland 2001), have been more frequently used in modeling, because the latter models can generate better predictions of subject reaction time in humans – especially in regard to the frequency of short-reaction time errors, and the shape of reaction time distributions in tasks with longer reaction times (Ratcliff and Smith 2004). A more recent accumulator variant, the Linear Ballistic Accumulator (LBA), resolves these concerns by dispensing with sequential sampling and within-trial noise (Brown and Heathcote 2005; Brown and Heathcote 2008), though partially at the expense of its relevance to implementation in nervous systems.

An additional parameter specifying an exponential decay coefficient for the decision variables is often added to race models, to simulate imperfect memory (Usher and McClelland 2001). The effect of this parameter is that evidence early in the stimulus time course is weighted less heavily in computing choice than recent evidence. However, a recent study of rat and human two choice decision making with a fixed stimulus duration provided evidence that memory for early evidence can be nearly perfect (Brunton, Botvinick et al. 2013), obviating the need for this parameter in at least some conditions.

1.4.2 Drift diffusion models

The drift diffusion class of choice models has arguably been favored in recent years, for their improved ability to model choice and reaction time across a broad range of discrimination tasks (Bogacz, Brown et al. 2006; Ratcliff and McKoon 2008; Resulaj, Kiani et al. 2009; Milosavljevic, Malmaud et al. 2010; Pleskac and Busemeyer 2010; Brunton and Brody 2013). In the drift diffusion model for a choice between two options, a single decision variable is continuously updated with signed evidence; evidence for the first choice adds value, and evidence for the second (null) choice subtracts value. Gaussian noise is added to the decision variable at each time step. When the absolute value of the decision variable exceeds a fixed threshold, the sign of the decision variable determines the choice (Ratcliff 1978). This single decision variable shorthand is equivalent to the case of two separate decision variables racing towards a fixed threshold, with perfectly anti-correlated evidence and noise (example trial shown in panel F of figure 1.3). Additional parameters are commonly added to the classic drift diffusion model to improve fit – most often two free parameters specifying a Gaussian decision variable starting point, to model choice bias or add response variability (Ratcliff and Rouder 1998).

To account for firing of choice-selective superior colliculus neurons in a monkey visual discrimination task, Ratcliff et al has developed a *dual diffusion* model of choice (Ratcliff, Hasegawa et al. 2007). In this variant of drift diffusion, two decision variables are used (to model separate,

competing populations of neurons accumulating evidence). Strength of evidence is scalar, and modeled as drift rate positively affecting one variable while negatively affecting the other by an equivalent amount, dispensing with the probabilistic component of sequential sampling of *evidence*. However, noise is modeled in time, and noise sources for two decision variables are independent. Which specific modifications to the classic drift diffusion model will ultimately explain the neural basis of decision making more generally – or whether a new class of models will be necessary, remains to be determined.

While classic drift diffusion models have been preferred to fit chronometric data in a wider array of studies than race models, Drugowitsch et al 2012 has shown that in a reaction time version of random dot motion discrimination, even better chronometric fits can be attained by modeling the temporal cost of continued sampling using a stopping rule with a collapsing threshold (Drugowitsch, Moreno-Bote et al. 2012) – an extension of the drift diffusion framework that had earlier been used to model urgency in monkeys (Churchland, Kiani et al. 2008). Apart from the drift diffusion framework, a similar collapsing threshold was learned from task contingencies by a reinforcement learning algorithm in an earlier study by Rajesh Rao (Rao 2010). A partially observable Markov decision process was implemented in an artificial neural network, and trained using temporal difference learning to perform the random dots task. One of the behavioral strategies the network learned was to adapt its decision making threshold as a function of sampling time, but only when the learning set was comprised of decisions under a deadline. It is not unreasonable that a decision maker interested in maximizing reward rate will strategically economize sampling time, much as economizing a different part of a trial (time investment) benefits decision makers in our temporal wager tasks (Section 1.3.4). However, introducing additional free parameters associated with constructing and using a cost function for time would add considerable complexity to our model of choice and confidence (Chapter 5) and we show that reasonable fits in most projections of choice, confidence and reaction time can be attained with a fixed boundary drift diffusion model. In section 5.7, we apply a simple collapsing boundary to improve our model’s chronometric fit for one subject.

1.5 Models of perceptual discrimination and confidence

In the beginning of this chapter, I invoked a Bayesian definition of decision confidence as the probability that a decision maker is correct, given the evidence used to decide (Kahneman and Tversky 1972). In our case, the evidence is a pair of Poisson click trains, which cannot be directly plugged into our definition and solved with Bayes rule. First, the click trains must be reduced to a measure of the amount by which the magnitude of chosen evidence exceeds the next best alternative as viewed by the decision maker, and this measure depends on how evidence is processed. The previous section introduced sequential sampling models of evidence processing in decision making. Several variants of these models have been extended to provide confidence reports. Among these extensions, the most straightforward applies to both the Vickers race model and the coupled accumulator model (a model that has a stopping rule with mixed absolute and relative characteristics). The confidence measure is simple and intuitive; when one of the accumulators reaches threshold, the difference between the winning and losing accumulator values is a measure of confidence (Vickers 1979). However, drift diffusion models cannot leverage the same measure; since the stopping rule is a fixed *difference* between decision variables, every trial would yield the same confidence value. Three alternative extensions have been proposed to provide confidence reports in drift diffusion: 1. to derive confidence from reaction time (Audley 1960; Ratcliff 1978), 2. to derive confidence from temporal variability in stimulus strength (Yeung and Summerfield 2012), and 3. to commit to a choice and sample *additional evidence* to determine confidence (Audley 1960; Pleskac and Bussemeyer 2010). Example trials of high and low confidence illustrating each of these measures are drawn in figure 1.2. In the remainder of section 1.5, I will introduce each model, and review how they fit the evidence they were initially intended to explain. In section 4.1, we show that each of these models fails to explain a set of critical patterns in our dataset.

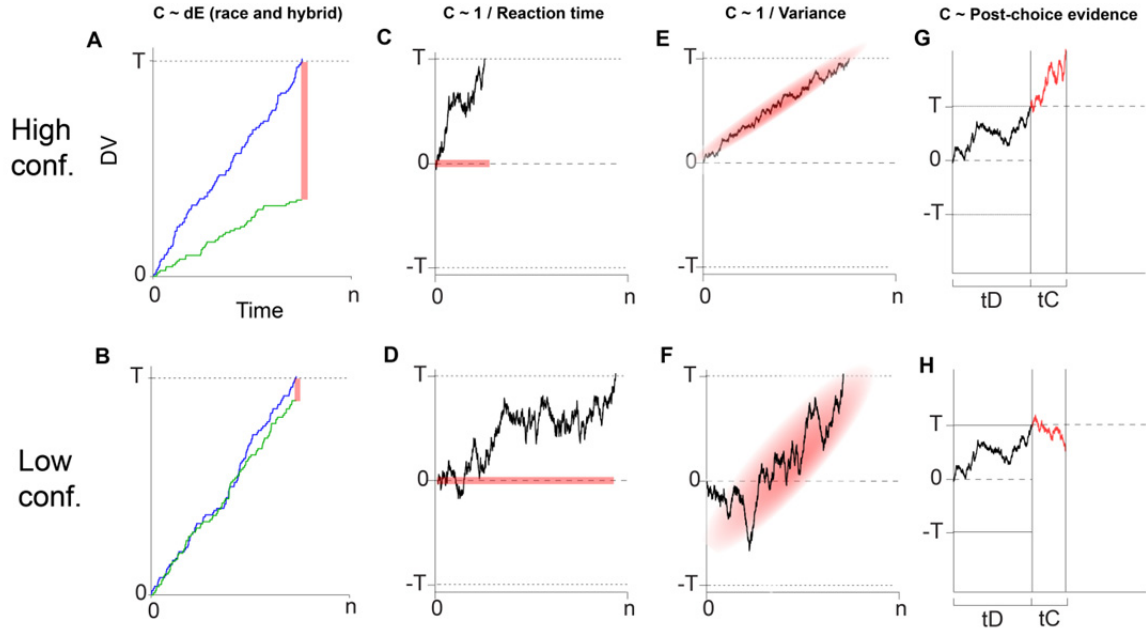


Figure 1.2: Measures of confidence in sequential sampling models of choice. For each measure of confidence, the time course of decision variables is drawn for one high confidence trial (top row of panels) and one low confidence trial (bottom row). The dotted line at T and/or $-T$ indicates the decision threshold. The decision variable properties used to compute confidence are indicated in red. **A,B:** In the race and coupled accumulator models, confidence is computed from the difference between winning and losing decision variable values. **C,D:** The reaction time model computes confidence from the total time taken to decide. Shorter decisions produce higher confidence. **E,F:** In the variance model, consistent evidence during the decision process generates high confidence, while irregular evidence generates low confidence. **G,H:** The post-decisional evidence model proposes commitment to a decision at threshold during variable time interval tD , followed by an additional phase of evidence collection. The agreement between the value of the post-decision variable and the original choice after a fixed time period (tC) determines confidence.

1.5.1 Accumulator model

In the sixth chapter of his 1979 book “Decision Processes in Visual Perception”, Vickers et al. proposes that confidence can be explained as the balance of evidence at the moment of choice in a fixed boundary accumulator model. Confidence is simply computed by equation 1.3,

$$C = E_C - E_{AC} \quad (1.4)$$

Where E_C is the value of the decision variable accumulating evidence for the chosen hypothesis, and E_{AC} is the same for the anti-chosen hypothesis.

Without providing fits to data, his argument focuses on four key predictions of his model that had been observed in prior studies: 1. the relationship between confidence and strength of evidence is positive and monotonic though shallower than the psychometric function, 2. The relationship between confidence and accuracy is nearly linear, 3. Confidence on correct trials is higher than confidence on errors, and 4. Confidence is inversely correlated with reaction time. In a subsequent study, Vickers provides evidence qualitatively consistent with assertions 3 and 4 in both speed and accuracy conditions, using a two choice line length discrimination task with a post-decisional confidence report on a 5 point scale (Vickers and Packer 1982). However, a model fit to data is not shown.

A related model – the Poisson race model (section 1.4.1), also uses a fixed stopping rule and balance of evidence to compute confidence, but processes evidence as discrete units in continuous time instead of continuous evidence in discrete time-steps. Merkle and Van Zandt et al (2006) tested subjects in both “easy” and “hard” conditions, where each condition consists of a 450 trial block of only easy or difficult trials. Subjects determined whether a screen contained more or fewer points than a mental standard, and reported confidence at the same moment as choice, on a bipolar 10-point response scale (containing 5 confidence divisions for each hypothesis and no neutral option). The authors show that this variant of the model can fit confidence calibration functions under both easy and hard conditions (Merkle and Van Zandt 2006).

In our own data, the race model can simultaneously fit psychometric and confidence functions (confidence with respect to accuracy and discriminability). However, in chapter 5 we analyze reverse correlations of the time-course of evidence triggered on choice, and show that the race model predicts a fundamentally different reverse correlation than what is apparent in our human subjects.

1.5.2 Coupled accumulator model

While the drift diffusion class of models is arguably favored over race models in decision making literature, it is unable to leverage the balance of evidence measure to compute confidence.

However, an intermediate class of models termed “coupled accumulator” or “hybrid” models improve upon classic drift diffusion chronometric fits (Usher and McClelland 2001) while retaining some usefulness of the balance of evidence confidence measure (Moreno-Bote 2010). The coupled accumulator confidence model functions as a race model, in that two stochastic integrators accumulate evidence towards a fixed boundary as in equation set 1.3. However, the two decision variables are correlated with a coupling constant ρ , such that evidence and noise arriving at one is partially subtracted from its opponent. This model is given as:

$$\begin{aligned} L_{i+1} &= L_i + E_{L,i} + \eta_{L,i} - \rho(E_{R,i} + \eta_{R,i}); \\ R_{i+1} &= R_i + E_{R,i} + \eta_{R,i} - \rho(E_{L,i} + \eta_{L,i}); \\ \max(L_i, R_i) &< T \end{aligned} \quad (1.5)$$

where E_L and E_R are vectors of evidence for left and right hypotheses at each time point, η_L and η_R are noise vectors drawn independently from a standard Gaussian distribution with a free variance parameter, and T is the choice threshold and ρ is the coupling coefficient. When $\rho = 0$, the model is equivalent to a race model, and when $\rho = 1$, it is equivalent to a drift diffusion model as illustrated in figure 1.3.

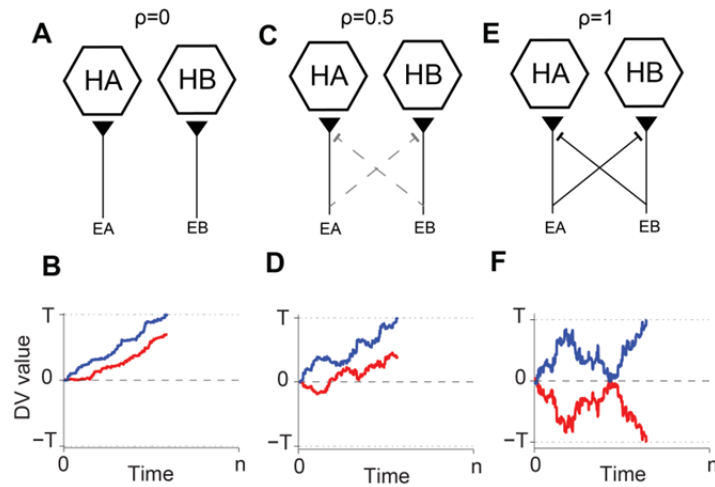


Figure 1.3: Processing of evidence in coupled accumulator models. Upper panels show a schematic of how two accumulators supporting hypotheses HA and HB use evidence from streams EA and EB. In lower panels, the time course of the winning decision variable is plotted in blue, and the losing variable in red. The dotted line at T and/or $-T$ indicates the decision threshold. **A,B:** The

race/accumulator model ($p=0$). Decision variables integrate evidence streams independently. The losing accumulator is a pure measure of evidence collected against the winning hypothesis. **C,D:** A coupled accumulator model ($p=0.5$). Lateral inhibition between the evidence streams diminishes the degree to which the final position of the losing decision variable represents the magnitude of evidence in stream EB (note the partial symmetry). **E,F:** With a purely relative stopping rule as in drift diffusion models, the losing decision variable is always the inverse of the winner, and its final position contains no information about the magnitude of losing-stream evidence.

The coupled accumulator model was tested in two discrimination tasks by Zylberberg et al, with emphasis on explaining the time course of evidence (Zylberberg, Barttfeld et al. 2012). Following initial experiments using a fixed duration random dot motion stimulus, the authors reproduced their results using a two choice luminance discrimination task similar to our Poisson click task (section 2.1). In the luminance discrimination task, subjects viewed two patches whose luminance was sampled every 40ms from Gaussian distributions with equal variance and different means. Subjects were free to respond by lateral eye movement once they determined which patch was brighter, and provided a post-decision confidence report with an eye movement to the desired position on a continuous confidence scale. The authors aligned chosen and anti-chosen stimuli to stimulus onset, and showed that the chosen patch was brighter than average, while the anti-chosen patch was dimmer than average, though this effect returned to baseline within the first 250ms of sampling. The authors then separated both time series by high and low confidence, and compute a “confidence kernel” for each evidence stream as the difference between high and low confidence time series. This analysis appears to show that only evidence from the chosen side is used to compute confidence. They then fit the choice and confidence “kernels” with race, coupled accumulator and drift diffusion models, determining that a coupled accumulator model is the best fit.

Their fits yielded a surprising result – the race and drift diffusion “choice kernels” looked identical (see our conflicting result in section 5.4). We suspect that the effect in their study may have been attributable to a methodological error – that the stimulus-aligned averages contained stimulus data that was never observed by the subjects, since the authors did not crop the stimulus-aligned vectors to response time. A choice-aligned re-analysis (provided by the authors in personal

communication) yielded a symmetrical reverse correlation when aligned to choice as in our study, indicating either a drift diffusion process, or a hybrid process with a very high coupling coefficient. The authors do not show model fits to psychometric or chronometric functions, or to confidence reports in any projection. We tested the fitness of balance of evidence in the race/coupled accumulator/drift diffusion spectrum in section 5.6, and show that these models can reproduce either the correct reverse correlation or the correct slope of errors in confidence v. discriminability, but not both.

1.5.3 Reaction time model

The reaction time (RT) model of confidence for drift diffusion was originally proposed by R. Ratcliff in his 1978 Theory of Memory Retrieval (Ratcliff 1978). This theory models a memory item's match or mismatch to a sensory probe as a drift diffusion process, and was subsequently adapted to explain two choice perceptual discrimination (Ratcliff and Rouder 1998). To test how well this theory fits human subject "familiar" or "novel" classifications, reaction times and confidence judgments, the authors used a recognition memory procedure. Four subjects were required to memorize lists of 16 words, and determine whether or not words had been shown previously, in a later probe session interleaving the 16 familiar words with 16 novel words. Subjects entered their "familiar" or "novel" classifications on a 6 point confidence scale, with three divisions for "familiar" and three for "novel" – similar to a Lickert scale though without a neutral option. Four difficulty conditions were created by adjusting how long the subjects had to memorize each list item; 0.5, 1, 1.5 or 2 seconds. The authors show that classification, reaction time and discriminability can be fit qualitatively by a drift diffusion model. In describing what predictions drift diffusion makes for these confidence ratings, Ratcliff notes that *"In terms of the theory presented here, the only information available about the "strength" of the item during the recognition processes is the comparison time. Thus, if a comparison is taking a long time, the subject may reduce his confidence and respond on a lower confidence key."* (Ratcliff 1978). While confidence reports were found to be correlated with

reaction time and a “schematic” of the model prediction is provided, a confidence v. reaction time fit of subject data to the drift diffusion model is not provided; nor is the relationship of confidence to accuracy or difficulty shown for either the subject or the model. The decision variable time course on individual high and low confidence trials with this measure were illustrated in figure 2.1c-d.

While Ratcliff was the first to propose that confidence arises as a result of direct interrogation of the subject’s perceptual estimate of their reaction time, a similar view had been suggested earlier by R.J. Audley (Audley 1960) – that confidence may be determined by counting the number of “vacillations” in a stream of perceptual evidence. For instance, a subject who experienced “ABAAA” would be more confident than a subject who experienced “ABABAABAA”, where A and B are units of evidence for each of two hypotheses. If a relative stopping rule is used and evidence arrives in discrete pulses as in Audley’s model, this is equivalent to the reaction time model – though it suggests that the time (and thus confidence) estimates are derived by observing the evidence itself rather than a separate mental judgment of elapsed time.

Several ensuing studies (reviewed in section 1.1.3) have established conditions under which confidence and reaction time are dissociable. Nonetheless, the RT model of confidence has been debated in subsequent reports, (Usher and Zakay 1993; Vickers 2001; Wilimzig, Tsuchiya et al. 2008) with none featuring an argument that conclusively rules out reaction time. We show in chapter 5 that in our data, confidence is also correlated with inverse reaction time for all subjects. However, we also show that inverse reaction time in a drift diffusion model is *not* anti-correlated with discriminability on error trials, a robust pattern we observe in human and rodent confidence reports, in our normative model and in statistical confidence.

1.5.4 Variance model

As an alternative way to acquire a confidence measure despite use of a relative stopping rule, Yeung and Summerfield proposed a Bayesian model of decision confidence, posited in sequential sampling terms but independent of the drift diffusion model. In this model, a decision maker actively

samples evidence and forms Bayesian posterior estimates of the probability that each possible hypothesis is supported by the experienced evidence stream until some stopping rule terminates the process, while in parallel, updating an internal measure of the variance of evidence in the stimulus (Yeung and Summerfield 2012). The variance measure then directly translates to decision confidence, while the greatest posterior determines choice.

While the variance model was not explicitly tested by its authors, precise control over the time course of stimulus information in our task affords an excellent opportunity to determine whether variance in provided evidence can account for confidence. In chapter 5, we explored this possibility, showing that the variance of click intervals between individual stimuli is not correlated with confidence in a consistent way. We also tested the variance of decision variable position along its course to threshold in the drift diffusion model, and found similar results.

1.5.5 Two stage dynamic signal detection (2DSD) model

Previous models of choice and confidence based on drift diffusion had attempted to extract a confidence measure from properties of the trial that are not direct measures of evidence strength, such as experienced stimulus variance or reaction time. The two-stage Dynamic Signal Detection model (2DSD) revives an alternative way to acquire a direct measure of stimulus strength from a drift diffusion process (Pleskac and Busemeyer 2010). The central proposal is that once the decision variable crosses a boundary, a second stage of evidence collection is commenced (illustrated in red, in figure 1.2g-h). During this second stage, evidence either supports or contradicts the initial choice, and the magnitude of agreement determines confidence. Since evidence in this phase is not terminated with a *relative* stopping rule, this relative evidence can reflect the statistics of the stimulus. Deriving confidence from post-decisional evidence as an alternative to reaction time in drift diffusion models had been suggested half a century earlier by Audley et al, though a specific model was not provided or tested (Audley 1960).

The authors tested how well the 2DSD model fits human data using two tasks; a perceptual task (line length discrimination) and a general knowledge task (city population comparison). Each task was separately tested under speed and accuracy instructions. The model was fit to subject choice, reaction time and confidence simultaneously using a modification of the QMP estimation method (Heathcote, Brown et al. 2002), originally developed to fit reaction time distributions. The best fit confidence calibration functions produced by the model provide reasonable fits to data in the accuracy condition, however data points in the speed condition on both tasks are almost all outside of 95% confidence intervals for subject data (Pleskac and Busemeyer et al 2010, figure 3). The authors report reasonable fits for the relationship between confidence and reaction time, but do not investigate other properties of confidence, such as the relationship between confidence and discriminability on error trials, or the time-course of patterned evidence contribution to confidence.

The second stage of the 2DSD model was explored by comparing two stopping rules; a fixed duration rule (used for the fits described previously), and an “optional stopping model” approximated with a Markov chain, whereby stopping is probabilistic and based on evidence strength. The fixed duration rule makes a strong prediction; information in the final part of a choice-aligned stimulus will be highly correlated with confidence and not with choice, while information earlier in the stimulus will correlate with choice and not confidence. We examine this prediction in chapter 5, and show that humans gather information for confidence from early moments of choice, while post-decision evidence is not as informative as predicted by 2DSD with respect to evidence collected during choice.

1.6 Head-fixed behaviors for studies of perceptual discrimination in rodent

Guided by insight about the abstract computation and the algorithm which the brain performs when determining confidence, we hoped to advance our understanding of the circuit mechanism underlying confidence in the brain. An array of emerging molecular tools hold great promise for the neuroscience of decision making, in their ability to isolate and manipulate elements of neural circuits. However, these tools are most easily used in a head-fixed mouse assay, a preparation that lacks a

precedent for the advanced multiple choice behaviors available in monkey and rat. In this section, we argue for the ideal of developing behaviors for head-fixed mice that are as complex and informative as freely moving behaviors previously available in rat. In chapter 6, we introduce a novel response device and behavior that constitute a step towards this ideal.

1.6.1 Advantages and disadvantages of the mouse model for behavior

In order to probe the neural circuits underlying behavior, the mouse has emerged as an ideal model system due to its genetic flexibility. Mice can be engineered to express fluorescent reporters of neural activity in genetically targeted circuit elements (Li, Burrone et al. 2005; Zariwala, Borghuis et al. 2012). In addition, cre-driver mouse lines targeting specific cell-types (Lindeberg, Usoskin et al. 2004; Tanahira, Higo et al. 2009; Taniguchi, He et al. 2011; Zariwala, Borghuis et al. 2012) have made a growing array of circuit components genetically targetable for studies seeking to examine their respective roles in neural computation. In recent years, experiments combining use of imaging techniques with behavior in mice have exploited these advantages using various head-fixed preparations (Dombeck, Khabbaz et al. 2007; Andermann, Kerlin et al. 2010; Komiyama, Sato et al. 2010). Studies conducted in awake head-fixed mice that require a behavioral choice readout have previously been limited to Go/No-Go behavioral responses (Mehta, Whitmer et al. 2007; Andermann, Kerlin et al. 2010; Komiyama, Sato et al. 2010; Histed, Carvalho et al. 2012) or virtual navigation (Harvey, Collman et al. 2009). Although considerable progress has been made using these classes of behavior in the rodent, epochs of disengagement and impulsivity in Go/No-Go tasks complicate the interpretation behavioral choices (Stuttgen and Schwarz 2008; Schwarz, Hentschke et al. 2010).

1.6.2 Go/NoGo and 2AFC task designs

While previously unavailable for head fixed mouse, two alternative forced choice (2AFC) behaviors have been favored in the neuroscience of decision making. The 2AFC design permits more informative interpretation of incorrect choices. In Go/No-go tasks, impulsive response behavior

appears as a bias towards the “Go” choice, while impulsivity in a 2AFC discrimination task can cause performance to decrease without affecting the readout of choice preference. Likewise, epochs of relative disengagement in Go/No-Go tasks are difficult to differentiate from a preference for No-Go responses. For these reasons, the process of making an active choice response on each trial makes behavior in 2AFC tasks considerably easier to interpret (Schwarz, Hentschke et al. 2010). Moreover, the symmetric reward contingencies of 2AFC tasks enable reaction time measurements that are not confounded by choice-specific motivational factors typical in Go/No-Go task designs (Zariwala et al. 2006). Indeed, a growing number of studies in freely moving rodent have taken advantage of 2AFC tasks using choice ports for choice readout with high temporal resolution. Rat 2AFC tasks using ports have been used to study the neural correlates of mental variables in decision making (Kepecs, Uchida et al. 2008), sensory processing (Otazu, Tai et al. 2009; Cury and Uchida 2010; Yoshida and Katz 2011) and motor control (Felsen and Mainen 2008; Erlich, Bialek et al. 2011). Similar 2AFC tasks have also been developed for *freely moving* mice (Rinberg, Koulakov et al. 2006; Busse, Ayaz et al. 2011). However, due to the small size of the mouse, head-fixed behavior greatly facilitates the use of contemporary techniques such as brain imaging, awake patch electrode recordings, high channel count microdrives and optogenetics.

1.7 Outline of chapters

Chapter 2. This chapter describes the methodology employed, especially for cases where a technique was relevant across chapters. I describe our auditory stimuli and other behavioral task contingencies, training protocols, technologies developed in support of research, data analysis and statistics.

Chapter 3. We examined the relationship between the *feeling* of confidence reported by human subjects, and *statistical* confidence. While relationships between decision confidence and other decision measures have been described, there is no clear set of properties that define a measure of confidence a priori, which could be used to interpret a behavioral or neural confidence signal. First

we derived signature interrelations between confidence and other decision variables from a normative definition of confidence – the Bayesian posterior probability of a choice, given the evidence used to decide. We then showed that Frequentist statistical confidence values (p-values) robustly show all three of these patterns in several statistical tests and noise regimes. Based on these results, we were able to show that the *feeling* of confidence reported by our subjects exhibits the signature patterns of a *statistical* confidence value. This result places confidence on a rigorous computational footing which is a necessary bridge from abstract variables to algorithms, by identifying the class of computation that generates confidence in the human brain.

Chapter 4. In order to understand the *mechanism* of confidence computation in the brain, Neuroscience research must rely on animal models. However, animals lack the linguistic ability to report confidence, forcing researchers to employ implicit measures that are captured as part of decision making behaviors. Investment decisions benefit from calibrated confidence, and can be precisely measured. We developed a time investment task which could be performed by humans and rats, to establish the relevance of an implicit confidence measure to both the human feeling of confidence and statistical confidence. This task was unique in allowing direct within-trial comparison of implicit and explicit confidence in humans. We determined that time investment in both humans and rodents strongly exhibited the signature patterns which we formally ascribed to statistical confidence in the previous chapter. We also determined that on a trial by trial basis, time investment was strongly correlated with verbal confidence in humans. Taken together, these results establish time investment as a behavior derived from confidence, forming an empirical basis for understanding confidence in rodents.

Chapter 5: Several conflicting models of choice and confidence were reviewed previously in chapter 1, which make differing predictions about how evidence is processed. To determine which *algorithm* is used by the human brain to classify evidence and compute confidence, we designed a new

confidence reporting task where choices were made based on high resolution temporally structured evidence – acoustically isolated streams of Poisson clicks. Depending on which algorithm the subject used to process evidence, different patterns in the temporal structure of clicks preceding each choice and confidence report should be evident. We tested five previous algorithms, and determined that none of them could simultaneously account for robust patterns in human confidence and in the history of evidence. To address this, we developed a new model – the Opportunistic Coupled Accumulator (OCA). We showed that OCA was capable of *simultaneously* approximating choice, reaction time, confidence, discriminability and evidence history data in several key projections for which previous models made qualitatively incorrect predictions. By describing a treatment of evidence on a moment-to-moment basis that is consistent with choice and confidence measures, this result provides crucial insight for studies investigating the circuit mechanisms that compute confidence in the brain.

Chapter 6. Genetically encoded molecular imaging tools hold great promise for functional dissection of neural circuit mechanisms contributing to behavior - and are most easily implemented in the genetically flexible head-fixed mouse. However, the go/no-go decision making behaviors previously available in head-fixed mice lack several key advantages of the two-choice behaviors available in freely moving rodents and monkeys, which make interpretation of behavior on individual trials more meaningful. To bridge this gap, we developed Choice Ball, a multiple choice response interface with which mice can report their decisions by sliding their front paws in opposing directions. Using this interface, we trained mice to make the same Poisson click stream classification decisions that were the basis of our decision confidence tasks, to determine how well this mode of response could capture the behavior of a deciding mouse.

Chapter 7. In the general discussion, I place the results of the individual chapters in the context of prior research, and discuss implications of our framework for future research.

Chapter 2

Methodology

Decision confidence has been studied for over a century, using a wide range of behavioral tasks, reporting scales and models. To determine which computational model best describes how the brain computes confidence, we developed a novel confidence reporting task that leverages high resolution, temporally structured sensory evidence to gain insight into how subjects compute confidence. A simple adaptation of this task enabled us to study confidence-guided time investment in rats, for the first time linking an implicit confidence measure available in multiple species to explicit confidence. In doing this, I developed two novel, open source instruments that will be useful in a wide range of behavior and stimulation applications. This chapter reviews the techniques used in the present confidence studies, with particular focus on novel methods. Methodology for the choice ball assay is introduced separately in Chapter 6.

2.1 Development of decision confidence reporting tasks for humans and rats

We had two primary goals in designing the tasks used in our study: 1. We sought to design a task that would provide insight into how evidence is used to generate choices and confidence reports. 2. We wanted to relate the human subjective notion of confidence to the time investment measure, and therefore constrained our task design to accommodate subjects of both species with appropriate provisions for capturing motor responses. We report on our realization of these goals in the remainder of this section.

2.1.1 The random click stimulus

An ideal stimulus for a decision making study contains attributes that give the experimenter insight into the decision process, and can be delivered precisely on each trial. Stimuli can be categorized as static or dynamic, depending upon whether the stimulus changes in a predetermined

way during sampling. Static stimuli have the disadvantage that subject errors are entirely the result of noise introduced in the perceptual, memory access and decision processes. How *specific instances* of categorically identical static stimuli were classified offers experimenters little insight into decision mechanisms that process evidence with respect to a choice threshold. Dynamic stimuli address this problem by providing evidence with a variable time course, generated by the experimenter. Looking back in time from the moment of choice, decisions can be explained by patterns in recent evidence, lending insight into how evidence is processed in time (Nienborg and Cumming 2009). In primate vision, the random dot motion (RDM) stimulus is a dynamic stimulus that has been used extensively to study the neural correlates of choice (Newsome, Britten et al. 1989; Roitman and Shadlen 2002; Gold and Shadlen 2007). Studies using RDM have taken advantage of the known time course of stimulus information (Adelson and Bergen 1985; Zylberberg, Barttfeld et al. 2012). We developed the random click stimulus as an auditory analog of RDM (Sanders and Kepecs 2012), while the same stimulus was independently developed by Brunton et al. (Brunton, Botvinick et al. 2013).

The random click stimulus, shown in figure 2.1, consisted of two independent streams of Poisson clicks, whose underlying rates were estimated and compared by the subject to solve each trial. Individual clicks were bipolar square pulses lasting $200\mu\text{s}$, and were presented to our human and rat subjects at $\sim 70\text{dB SPL}$, to ensure a high signal to perceptual noise ratio. Click streams were pre-computed for three-second trains by drawing click intervals from an exponential distribution whose mean was determined by the intended click rate. Click streams were truly Poisson; we did not force the click trains to contain the exact ratio of clicks needed for a perfect estimate of underlying rate. This allowed us to preserve the instantaneous expected rate for analyses where we averaged the stimulus across trials. The onset time and waveform information for each click in the stream pair was computed in MATLAB (Mathworks) and transmitted to the Pulse Pal device (section 2.1.6) prior to each trial. When triggered, Pulse Pal directly played the click train into speakers or headphones. In RDM tasks, subtle variation in a subject's visual fixation may affect how the subject perceived the intended stimulus on a given trial. However, click streams in our assay did not have the constraint of

mechanical visual fixation, and could be delivered reliably to the sensory periphery using a pair of ambient noise attenuating headphones (HD-280; Sennheiser). This level of stimulus control was not possible in untethered freely moving rat – instead, speakers (HP 5187-2105; Harman Kardon) were used on either side of the animal, presumably requiring an additional stream localization computation to assign each perceived click to a stream (Bregman 1994) in discriminating the higher rate. The difference between rates of the two streams was manipulated to adjust discriminability. For both rats and humans, underlying rates always summed to 100Hz, and precise rate pairs were determined by subject performance as described in section 2.1.3.

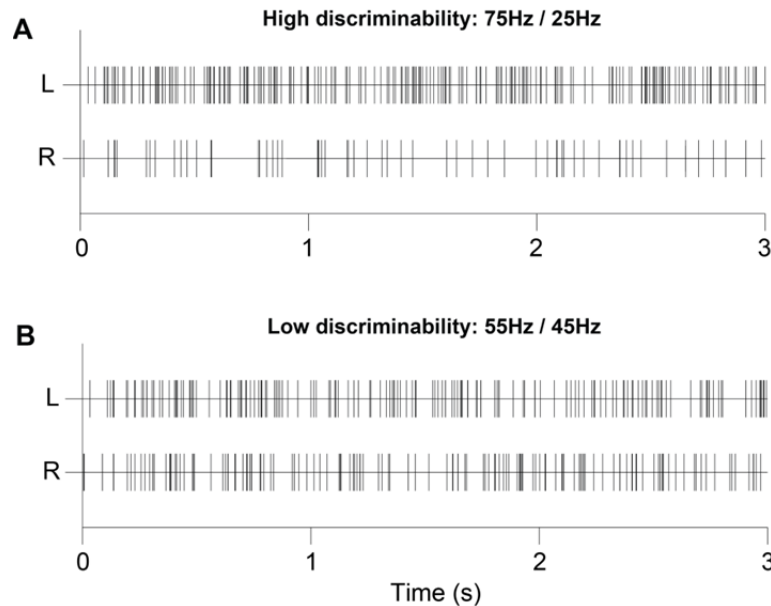


Figure 2.1: Time course of the random click stimulus. Subjects compared the rates of Poisson click streams delivered from left (L) and right (R) auditory channels. **A:** Example of a high discriminability (easy) stimulus, featuring a large difference between left and right channel rates. **B:** Example of a low discriminability (difficult) stimulus. Actual rates were adjusted to subject performance.

2.1.2 Human perceptual confidence task design

Subjects and compensation

Human subjects were recruited from the general population of Cold Spring Harbor Laboratory, using public notices. Subjects accepted into the study verbally agreed to participate for at

least eight 60 minute sessions, and signed a statement of informed consent. All subjects reported normal or corrected-to-normal vision and did not have hearing deficits to the best of their knowledge. Subjects were technicians or graduate students, including 2 males and 5 females, ranging from 23 to 30 years of age. Two of seven subjects were left handed. Five additional subjects participated, but none completed more than three sessions before voluntarily departing the study. One further subject was eliminated post-hoc for using a “first-click” detection strategy (section 3.1.2).

Subjects were paid \$0.07 for each correct response in the Explicit-only variant of the task (see below), and \$0.10 for each correct response in the Implicit variant. Prior to participation, subjects were informed that random guessing would earn them at least \$7 per hour if they sampled each entire stimulus, and that accurate responses could earn them in excess of \$20/hr. Such large incentives were necessary to encourage subjects to use time within trials optimally. Subjects were periodically paid in cash, when they requested a payment or when their balance exceeded \$200.

An online calendar (Google) was used to coordinate participation among subjects. Participants were free to schedule sessions at their own convenience, and were given unique 8-digit subject ID codes to log in and start the experiment. The calendar was monitored to ensure that subjects completed at least 2 sessions per week.

Apparatus

A single apparatus was available for all subjects, in a dark room with 24 hour lab ID cardholder access. A personal computer (Inspiron 660, Dell) powered the experiment using the Windows 7 operating system (Microsoft). Instrument control and data collection were accomplished in MATLAB r2011a (Mathworks), with two software extensions: the MATLAB statistics toolbox (Mathworks) and the Psychophysics toolbox (Brainard 1997). Commercial computer peripherals relevant to the experiment included a USB joystick (ST290, Saitek), and a standard USB keyboard (Dell) and an LCD monitor (G245HQ, Acer) positioned at eye level, ~70cm in front of the subject.

For confidence reporting, custom-printed adhesive labels were placed over keys on the keyboard's keypad as shown in Figure 2.2.



Figure 2.2: Layout of the confidence response keypad. Subjects entered confidence on a 5 point scale, which was presented vertically to capture intuition about high and low, and to discourage confidence-side associations in our sided response task. Subjects were instructed that “5” indicated high confidence, and “1” indicated a random guess. Since decision confidence is confidence in the hypothesis that a choice was correct, the “Err” button was provided to indicate that the subject was so uncertain that they believed their previous choice was likely to be wrong.

A real-time click generator and response capture device (Pulse Pal, section 2.1.6) interfaced with the computer by USB. Prior to each trial, it received the trial's initial delay, timeout and click times from the governing computer. During each trial, it returned button press and release events. Following each trial, it returned reaction time and a temporal wager (if applicable, Figure 2.3) each with 50μs resolution. In the explicit-only task variant, the time between the choice and confidence report was measured by the PC clock, using low level keyboard polling and timing functions in the psychophysics toolbox to attain maximal temporal precision. A pair of headphones (HD-280, Sennheiser) was connected to the real-time click generator's output channels with appropriate wire adapters, for stimulus delivery. A pair of PC speakers was connected to the computer's motherboard sound port, and positioned on either side of the subject beneath the monitor, for feedback sounds.

To encourage long term human participation and aid with recruitment, a custom appetitive reward system was installed above the computer. Three servo motors (HSR125-CR, Hitec) powered and controlled by a USB servo controller (Micro Maestro, Pololu) were governed with custom

software written in MATLAB. Each servo motor was connected to the crank of a confection dispenser (Mini Bean Machine, Jelly Belly USA), and was secured in place using wire ties. Plastic pipes converged the dispensed product from the dispensers, through a PVC slanted cross fitting (Lipson International) into a Tupperware container positioned on the subject's left side at arm level. The appetitive reward system was only used in the break periods between blocks of trials.

Data were automatically stored by the behavior computer on a network drive (Google) labeled by subject ID code to ensure anonymity, were the data to be compromised. The key relating subject ID codes to subject identities was encrypted, and stored on a removable drive by the experimenter. Data analysis was performed in MATLAB using custom scripts.

Procedure and training

Two variants of the human perceptual confidence task were designed. Figure 2.3 shows illustrations of the flow of events for each task variant.

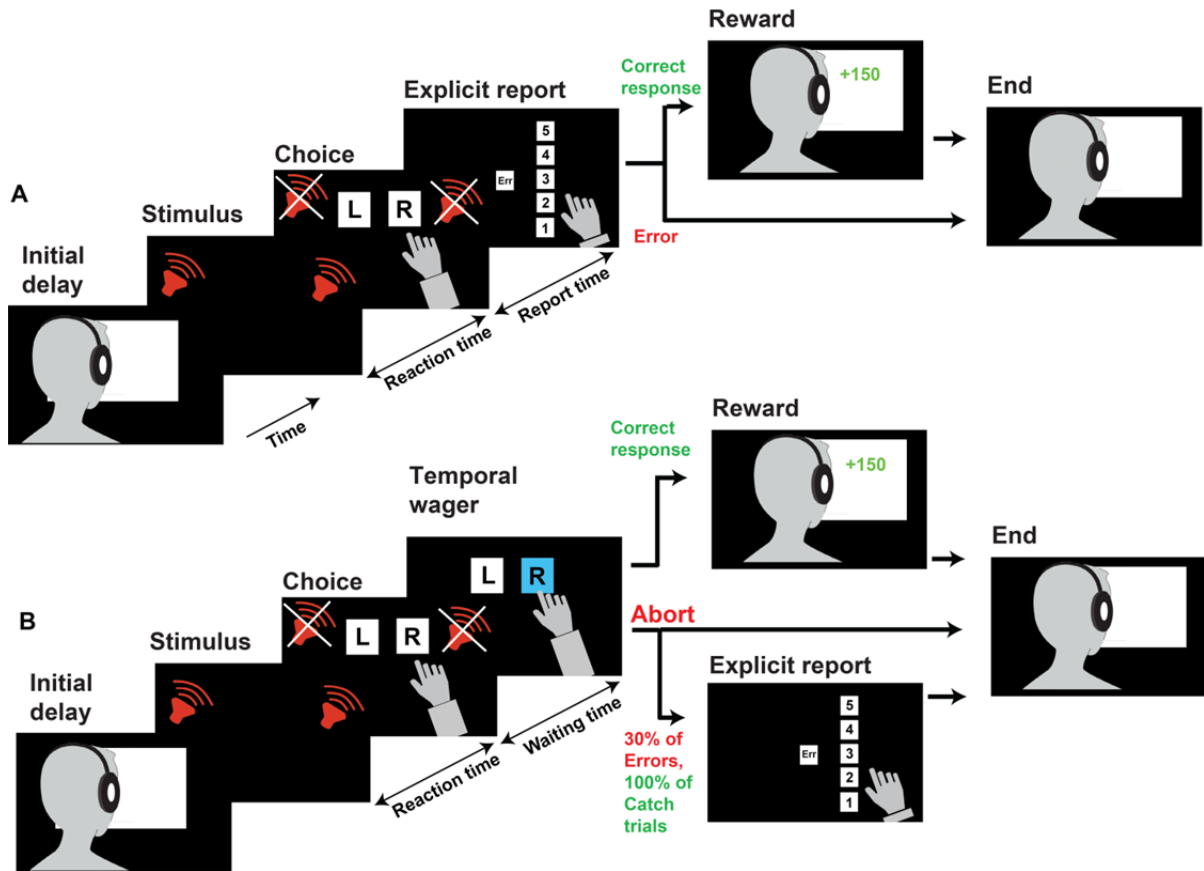


Figure 2.3: Flow of trial events for explicit-only and implicit tasks. Once a trial was initialized by the subject, the subsequent event flow through five types of states is shown: Initial delay, stimulus sampling, choice, confidence report and reward feedback. **A:** The explicit-only task. **B:** The implicit task.

The purpose of the explicit-only task variant was two-fold: to acquire explicit reports immediately after choice, and to acquire a large amount of data for in-depth analysis of confidence properties (since this task variant did not rely on omission trials at 10% prevalence). In the explicit-only variant, subjects were notified that each trial was ready with a chime played over the PC speakers, and they subsequently initialized each trial by squeezing the joystick trigger. On initialization, temporal control of the trial was delegated to modified firmware on the pulse pal device (section 2.1.5). Following an exponential random delay ($\mu=1.5s$), the click train was presented. Subjects listened to the click train, and entered their choice of which click train had a faster underlying rate (left or right) on the response device by pressing left or right response buttons. If

sampling time exceeded 3 seconds, the click train was automatically terminated, and the subject was informed visually that they had timed out. Within 50 μ s after a response button was depressed during sampling, the click train was terminated. Subjects were then visually prompted to enter a confidence rating. The visual prompt was a drawing of the red-labeled keys as shown in Figure 2.2, with two additional notations: “High confidence” above the “5” key, and “Random guess” below the “1” key. No time limit was imposed on confidence reports. When subjects entered their confidence, they were immediately informed if their choice was correct by a visual animation, accompanied by a “payoff” sound played over the PC speakers. Subjects were given no feedback if they erred. The total accumulated reward in the current block was displayed in the upper-right corner of the screen. Subjects entered their confidence following each choice on a labeled keypad (Figure 2.2).

Since decision confidence is confidence in the hypothesis that a choice was correct, the “Err” button was provided to indicate that the subject believed their previous choice was an error. None of our subjects predicted their errors with greater than chance probability, determined by a 95% binomial confidence interval evaluated on the outcomes of “Err” trials. Since subjects could not predict their errors, perceived error responses were counted as lowest confidence responses. In total, 190/17788 confidence reports were “Err”.

Contrary to the explicit-only task variant, the implicit task was a low-throughput assay for humans (since it acquired confidence reports on correct trials only 10% of the time). However, this task variant was very similar to the rodent task, and allowed implicit measures available across species to be directly compared with explicit reports. The Implicit task variant shared the same flow of events as the Explicit-only variant, with three changes: 1. Instead of an explicit confidence report after making a choice, subjects pressed and held their response key, waiting for the “correct” feedback to be presented. The reward delay was drawn from an exponential distribution ($\mu=1.25$ s) with a minimum cutoff at 0.5s to discourage impulsivity. If the subject was incorrect, the system would wait indefinitely until they released the choice button, indicating the end of their time investment in the expected reward. 2. Ten percent of trials were “catch trials”, on which reward and

reward feedback for correct decisions were omitted. On these trials, even if correct, subjects were eventually forced to abort their wait for reward, providing a measure of willingness to wait given a correct choice. How long a subject would wait before giving up is the measure referred to as “time investment”. 3. Subjects were prompted to enter explicit confidence reports following time investment, but before feedback on 30% of error trials and on all catch trials.

Three levels of stimulus difficulty were randomly interleaved, and a neutral evidence stimulus was added randomly at 10% prevalence with the constraint that two neutral evidence trials could not occur sequentially. In the first several sessions, the click ratios of the three difficulty levels were adjusted manually between sessions, to achieve target accuracies of 60%, 75% and 95%. Though initially a staircase procedure (Cornsweet 1962) had been implemented for this purpose in our pilot study, it was found to take longer than manual adjustment.

Subjects were first introduced to the apparatus by demonstration, and were given an instruction sheet explaining usage of the apparatus, which was available in the testing room for the remainder of the study. Subjects were encouraged to use the entire confidence scale, and advised that they may need to “re-map” their initial range of experienced confidence to span the scale divisions. Subjects were not given explicit speed or accuracy instructions. Rather, in the interest of similarity to the rodent task, they were instructed to earn as much payment as they could in the time allotted. The first two sessions were used for training and to adjust stimulus discriminability to subject performance. These sessions were not included in the remainder of the analysis. Following each of these training sessions, subjects were given written feedback about side bias, scale usage and performance. Once performance levels were within 5% of target accuracies (typically after 2 or 3 sessions), this type of feedback was no longer given. Subjects typically learned the instructions immediately, reached peak performance within two sessions, and did not substantially improve their performance for the remainder of the experiment (section 3.1.2).

Game interface rationale and implementation

In a pilot study, all but two of 9 human subjects had poorly calibrated implicit confidence, and often withdrew from the study after a few sessions. We attributed these shortcomings to several factors: 1. Since our analysis of the time investment measure relied on catch trials at 10% efficiency, we needed datasets of thousands of redundant trials. The students enrolled were often too busy or unmotivated to participate regularly. 2. Confidence is a measure of the quality of perceptual evidence, requiring sustained alertness during long sessions. Subjects reported difficulty remaining alert. 3. Economizing time investment provides an advantage in terms of rewards accrued. A water-restricted rodent from whose perspective water is an essential for survival, is supremely motivated to economize time investment. The same motivation is usually not true of human subjects by default.

To address these issues, I designed a video game surrounding 15 identical five-minute blocks of trials. Video game interfaces for psychophysics tasks have been used previously to encourage focus and participation in children (Abramov, Hainline et al. 1984; Soderquist and Shilling 1992). Since the subjects were comparing Poisson click streams, similar to the task of comparing distant Geiger counters to localize a radiation source, the game was titled “Plutonium Miner” on subject recruitment materials and consent forms.

On logging in to the study with their subject ID, subjects controlled a sprite avatar of a miner using a joystick, who could walk between mining sites and buildings on a game map (Figure 2.4a) where each site represented a block of trials. In order to access each site, the subject needed to purchase access at a “land office” building on the map. The land office provided a way to verbally remind subjects to gain as much reward as possible in the time allotted for each block, and to explicitly equate game currency with block-time. On entering a site, the subject committed to a five minute block of trials. Few visual animations were shown during the trial. During self-initiation and the initial delay, a sign post in a static field of random dots was shown (Figure 2.4b). The dot field was generated randomly on each trial. Once the initial delay had completed, the sign post disappeared and the subject’s avatar was shown. No visual changes occurred during sampling of evidence and

choice. In the explicit task, the choice was immediately followed by the confidence report prompt. In the implicit task, the avatar was shown facing the chosen direction at a fixed position in the center of the screen, while a field of random dots scrolled in the opposite direction. This animation was intended to give the sense that the avatar was actively exploring the chosen direction while the subject held the button down. If rewarded, the avatar would jump repeatedly and the reward amount was shown flashing above its head. For each correct trial, the subject earned 200 units of currency, and the cost to purchase access to the next block of trials was determined as $\frac{1}{4}$ of the subject's total profit from the previous block (though this ratio was not explicitly disclosed to the subject). Subjects could spend currency between blocks, by moving their character to a "business center" on the map, where they could construct M&M, Skittles or Reece's factories of varying efficiencies for fixed prices. During each inter-block map navigation time, subjects were verbally notified that each of the factories they had built had completed a quantity of product, which was then delivered to the subject by the appetitive reward apparatus. Subjects were given the choice to opt out of the appetitive reward delivery, though none chose to exercise this option.

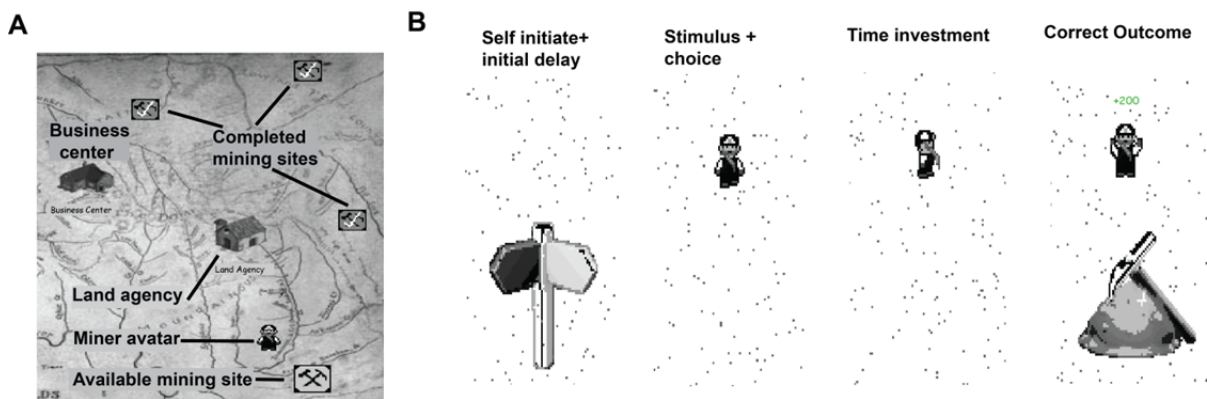


Figure 2.4: Game elements to encourage alertness and long-term participation. Subjects performed the task in blocks with identical trial parameters, lasting 5 minutes each. **A:** Completed blocks were represented on a "map" screen as dark pick-axe icons with white check marks. The next available block was shown as a lighter, slightly larger unchecked icon (near figure bottom). Subjects started each block by moving a miner avatar to the next available mining site, after purchasing rights to the site at the land agency (near map center) with in-game currency. Between blocks, currency could be spent at the business center (upper-left) to make purchases as indicated in the main text. **B:** Visualizations were shown during each trial. To prompt the subjects to self-initiate each trial with a

button press, a bidirectional sign post was shown on a randomly generated white field of black points (left-most panel). After a random delay, the subject was shown their avatar in a static field of dots while listening to click streams and choosing a direction. In the time investment task, the avatar was shown facing the chosen direction in the center of the screen while random dots moved in the opposite direction, to create the illusion of motion as the avatar searched for plutonium. If a trial was rewarded, after both implicit and explicit confidence measures had been acquired, a correct outcome was indicated with a two second sprite animation of the miner jumping alongside a pile of plutonium, while the static reward amount of 200 was indicated in green text above its head.

While I do not have comparative data to determine which (if any) of the game elements were responsible for improved subject retention and performance between our initial attempts at measuring confidence and the principal study, I took measures in the game design to minimize interaction with the decision making and confidence reporting processes. The only phase of a trial that is concurrent with changing visual game elements is the time investment period in the implicit study, during which the avatar is shown running against a field of randomly generated dots. All other dynamic game elements are shown between blocks, and ought to preferentially affect responses in the trials following a block interval. The subjects accrue virtual currency over the course of the session, much as our rats accrue relief from thirst. The stability of measures within sessions is investigated in section 3.1.7. However, unlike rodents, subjects *spend* currency between blocks to access subsequent blocks, and develop game assets. We show that mean performance, explicit confidence and time investment for blocks are uncorrelated with currency spent in the previous block interval (Figure 2.5).

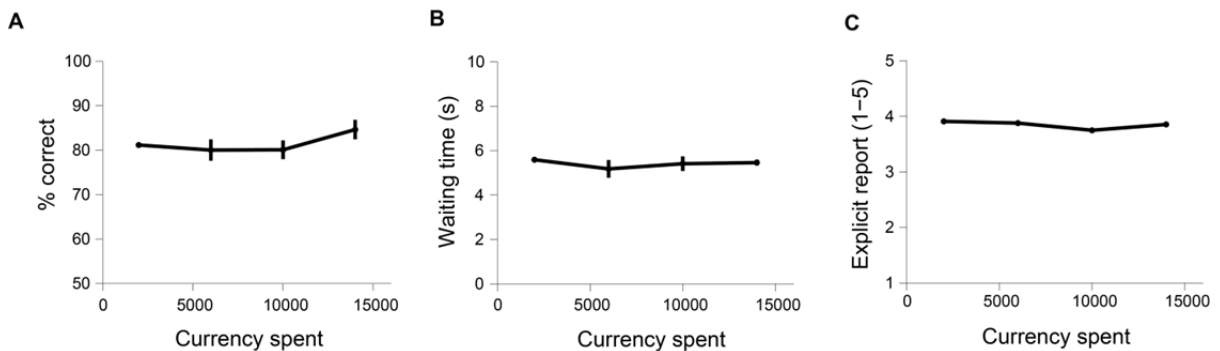


Figure 2.5: Spending game currency does not affect block mean performance or confidence. Data shown are means for 5 test subjects who performed both explicit-only and implicit tasks. The “currency spent” value for each trial denotes game currency spent in the inter-block interval prior to the current block. **A:** Choice accuracy as a function of currency spent in the prior block interval. **B:**

Time investment as a function of currency spent. **C:** Explicit confidence report as a function of currency spent. Errors show 95% CI of the mean.

2.1.3 Rodent perceptual confidence task design

Subjects

Data are reported from six male Long Evans hooded rats, aged 8 weeks at the onset of training. Rats were housed in pairs on a 12-h reversed light/dark schedule. Food was available ad libitum, and the rats were placed on a liquid restriction schedule with daily body weight monitoring to ensure that body mass remained within 85% of mass before restriction. For 30 minutes each day following training, rats were given free access to water. Two additional rats were used, but did not learn to suppress impulsivity as required to wait for delayed rewards at the apparatus response ports. These rats were excluded from the study after 1 month.

Apparatus

Sessions were conducted in a custom operant conditioning chamber, enclosed in a ventilated acoustic isolation chamber (Industrial Acoustics). Three walls of the operant chamber were featureless 40cmx40cm black acrylic panels. The fourth wall was fitted with three custom ports, containing infrared photo-gates to measure snout entry and exit. Sounds were delivered using a set of speakers (HP 5187-2105; Harman Kardon) positioned behind the acrylic panels adjacent to the response port wall. Speakers were calibrated to 70-dB sound pressure level (SPL) for the click stimulus using a pressure-field microphone (Brüel & Kjær, Sound & Vibration Measurement, Nærum, Denmark). Water reward was dispensed from reservoirs above the box into the outer two ports through silicone elastic tubing. Water flow was controlled using a solenoid pinch valve (NRResearch) with pulse timing calibrated such that a single reward measured 20µl. Behavior was governed by a PC (Island Motion Corp) running the Windows XP operating system (Microsoft) and MATLAB r2009b (Mathworks). An infrared camera affixed to the ceiling of the acoustic chamber

was used for observation of behavior on the PC. A top-down drawing of the port wall of the behavior box is shown in Figure 2.6.

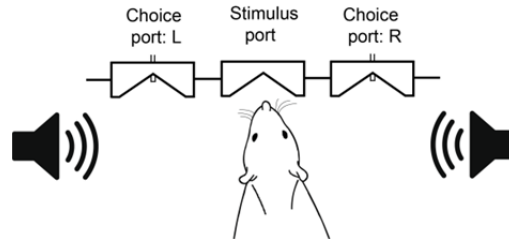


Figure 2.6: Rodent interface port configuration. Rats entered the central stimulus port to initialize each trial. Sound was delivered from speakers on either side of the rat until it withdrew from the stimulus port. The rat then provided its classification of the sound by entering the left or right choice port. If correct, 20 μ l of water was delivered from a drink tube in the choice port.

The experiment was programmed and implemented using a custom real-time behavior system I developed (Bpod, Appendix II). In brief, each trial in the experiment was programmed in MATLAB as a matrix of states and state transition conditions. This matrix was sent to the Bpod device, which executed the matrix in real-time (directly controlling lights and solenoid valves, triggering sound delivery and recording timestamps of port entry and exit events). During trials, event byte codes were sent to the PC as they occurred. When trials finished, timestamps of events were returned to the Bpod client on the computer, which saved data and displayed relevant statistics for online monitoring.

The Poisson click stimulus was delivered to the speakers by Pulse Pal (section 2.1.5), a triggered arbitrary waveform playback device I developed with support for parametric pulse stimuli commonly used in neuroscience. Pulse parameters were sent to Pulse Pal at the beginning of each session, and click times were sent prior to each trial via USB. Firmware changes from the public default were unnecessary since unlike the human variant of the task, the Bpod device measured behavioral responses and Pulse Pal was only used for sound delivery. Bpod triggered Pulse Pal by TTL pulse, over a BNC cable connecting the two devices. A subsequent TTL pulse delivered during stimulation stopped the ongoing stimulus within 50 μ s.

Procedure and training

The flow of trial events is shown in Figure 2.7. A trial was initialized when the rat broke the stimulus port photo-gate. Following a random delay (drawn from exponential distribution, $\mu=200\text{ms}$, re-drawn if $< 50\text{ms}$ or $> 500\text{ms}$), the click train was delivered from the speakers until the rat exited the stimulus port. It then had 3 seconds to enter a response port in order to continue the trial. On entering a response port, the rat waited for reward. Reward was delayed by an interval drawn from an exponential distribution ($\mu=2\text{s}$, re-drawn if $< 500\text{ms}$ or $> 8\text{s}$). If the rat exited the response port, it had a 500ms grace period to re-enter and continue its wait. 10% of trials were catch trials, on which reward was omitted even if the rat was correct. Sessions continued until the rat stopped performing trials by visual inspection, or until 2 hours had elapsed.

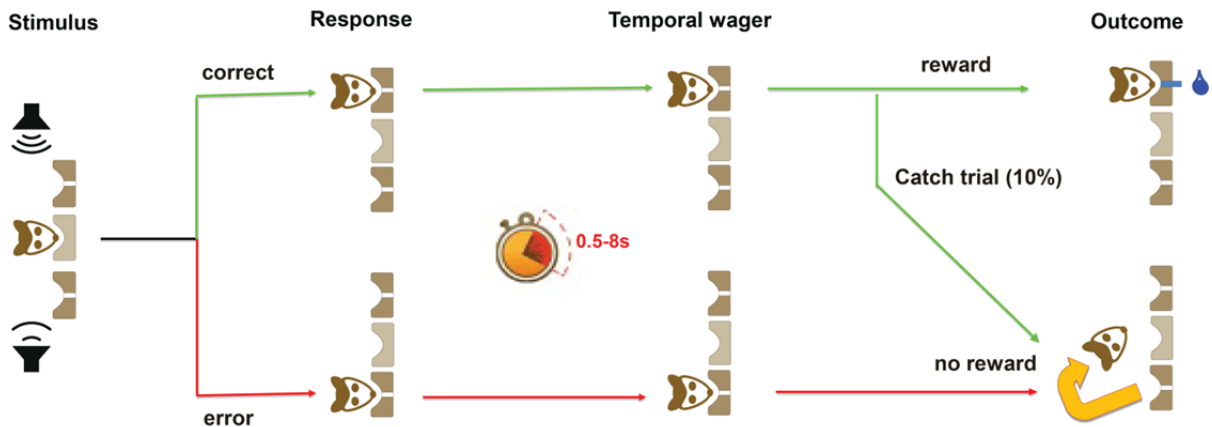


Figure 2.7: Flow of trial events in the rat implicit confidence task. Each trial can be considered to have four phases, each of which contained several events. In the stimulus phase, the animal initialized the trial by entering the stimulus port. Following a delay, the click train was delivered until the rat left the port. In the response phase, the rat entered one of the two response ports. In the temporal wager phase, the rat waited in the response port for reward. In the outcome phase the rat was either rewarded, or it aborted its attempt by leaving the port.

The training protocol included 5 phases: 1. Operant conditioning. 2. Sampling duration conditioning. 3. Simple two-choice discrimination. 4. Psychometric discrimination. 5. Adaptive reward-delay training. Each phase is described below. Training was conducted daily, and typically took 1 month to reach our training criterion.

Operant conditioning

Following one day on the liquid restriction schedule, naïve rats were introduced to the conditioning chamber after a brief handling period. The stimulus port was covered with masking tape, permitting them to explore only the response ports. Response ports delivered water on an intermittent reinforcement schedule when the rat entered either of them with its snout. The ratio of pokes to drops delivered proceeded as follows: for the first 50 rewards, the number of pokes required to elicit a drop increased by one poke every 10 trials (starting with 1 poke per drop at trial 1, 2 pokes per drop at trial 11, and so forth). Following trial 50, the poke ratio for each reward varied randomly until the rat was sated. If the rat did less than 200 trials before becoming disinterested, the Operant protocol was repeated on the subsequent day.

Sampling duration conditioning

Rats were placed in the conditioning chamber. No tape was placed on the stimulus port, allowing the rat to explore it for the first time. For the first 20 trials, a drop of water was delivered to each of the response ports automatically, at the moment the rat entered the stimulus port. Following trial 20, the rat had 5 seconds to enter a response port after exiting the stimulus port in order to get reward. Following trial 40, Poisson clicks were delivered at 50Hz from the speakers. The click trains did not instruct reward location, and could be ignored. Following trial 50, the minimum duration the rat had to remain in the stimulus port in order to be rewarded was increased from its starting point at 0ms, by 100ms every 50 rewarded trials. Rats that reached 500ms delays were advanced to the two choice discrimination task. Rats that did not were trained on this stage once each day until they met the criterion for advancement.

Simple two-choice discrimination

In this stage, the rat entered the stimulus port to initiate a trial. On each trial, a Poisson click stream was delivered from *either* the left or right side. The rat was required to remain in the stimulus port for 200ms before responding. If the rat exited before this minimum sampling time or classified the stimulus incorrectly, a 5 second punishment interval was enforced, during which the ports were non-responsive. The rat could only receive reward for entering the response port on the same side as the sound. When rats performed a single session with above 90% accuracy, they were advanced to the next stage of training.

Psychometric discrimination

The flow of trial events on psychometric discrimination was identical to two-choice discrimination, except that stimuli of graded discriminability were used. Rats were first introduced to 95Hz/5Hz stimuli. Once performance exceeded 90% for one session, the 95/5 stimuli were randomly interleaved with 75/25 stimuli with equal probability. Once performance exceeded 80% for one session, 65/35 stimuli were added with equal probability. Exact click rate ratios were then adjusted manually between sessions for each rat to create three difficulty categories: easy trials (95% classification accuracy), medium trials (75% accuracy) and difficult trials (60% accuracy). Since performance at these levels was not requisite for the subsequent stage of training, rats were advanced as soon as the 65/35 trials were added.

Adaptive reward-delay training

Rats performing psychometric discrimination were trained to wait for reward at the goal port. The governing computer automatically increased reward delay from 0ms to 6s in 200ms increments, incrementing each time the rat waited for long enough to be rewarded on 80% of the previous 50 correct trials. This was typically the longest phase of training, taking several weeks in most rats.

Rat auditory confidence task

Following training, the intended task was repeated daily for several months to measure behavior. Since rats had successfully learned that rewards could be delayed for long periods, we reduced the mean time investment to 2s and varied the reward delay by an exponential distribution. To prevent impulsive withdrawal, the minimum reward delay was set at 0.5s. To prevent unrealistically long waits, a maximum delay was set at 8 seconds. Sessions contained two new types of trial: neutral evidence trials (50Hz/50Hz) which were rewarded randomly, and catch trials which were not rewarded at all. Neutral evidence trials occurred at 5% prevalence, and catch trials at 10% prevalence. For several weeks, rats were tested with mean time investment set to 2 seconds. Following this period, we reduced mean time investment to 1.25 seconds as was true for our human subjects, and captured several more weeks of data.

2.1.4 A quiz to study the role of sensory pipeline noise in confidence

We wondered whether the patterns we observed in confidence reports in the Poisson click tasks would persist in a regime where all errors were attributable to *non-sensory* cognitive mechanisms. To this end, we designed a confidence reporting task where we could confirm that perceptual noise did not corrupt the signal used to make a choice. The task required subjects to determine the larger of two quantities that are general knowledge – the populations of countries, and provide an explicit report of their confidence in their choice.

Subjects and compensation

Since longitudinal data and high temporal precision were unnecessary to answer our question, I formulated the task as a MATLAB geography quiz that could be distributed by email to potential student participants in our department. The recruitment email explained that the quiz would take 15

minutes, that the results would be anonymous apart from age, gender and handedness, and that there was no monetary compensation. It also specified supported computing platforms and provided installation instructions. Finally, the recruitment email requested that subjects not deliberately study geography or demographics prior to taking the quiz, to maintain that the quantities being compared were general knowledge and not dependent on a recent learning event. In addition to the subjects recruited via email, two visitors of our lab requested to participate.

24 subjects including the author returned data. Subjects ranged from age 20 to 50, and included 15 males and 6 females. Three subjects were excluded from analysis for mean session performance below 60% accuracy. Only one subject reported left-handedness, and none were ambidextrous. Each subject viewed a PDF copy of the study's informed consent document, and was required to fill a checkbox indicating that they had read and understood it before proceeding.

Apparatus

Since the subjects used their own computers to run the task, it was critical to ensure that the software worked as intended on a wide variety of platforms. Short sessions of 10 trials were run on the following operating systems prior to distribution: Windows 7, Windows Vista, Windows XP, OSX 10.6 and 10.7 and Ubuntu 12.10. The software was also validated on MATLAB r2009b, r2010a, r2011a and r2012a, though not every possible combination of platforms and software was tested prior to release. Reaction time measurements were made with the system clock command in MATLAB (*now*), and thus vary with respect to the moment when the word pairs were made visible to the subject. Uncontrolled factors were the monitor refresh rate, the program loop execution speed on different systems, and variable processing jitter determined by the host's operating system. We report trends in reaction time which occurred on the order of seconds, and were averaged across hundreds of trials.

Subjects controlled the task with the computer keyboard. The task instruction screen presented to the subjects is shown in figure 2.8.



Figure 2.8: Confography subject instruction screen. Subjects were asked to compare the populations of two countries within three seconds and provide a post-decision confidence report.

Country population data for the year 2012 was downloaded from the World Bank’s Open Data database at <http://data.worldbank.org/indicator/SP.POP.TOTL> and modified before inclusion in the quiz data file. To produce a continuous range of discriminability, countries with greater than 500 million or less than 10 million people were excluded. To reduce sensory processing time, I manually replaced countries with long official names with my estimation of their vernacular names (for instance, “Arab Republic of Egypt” was replaced with “Egypt” and “Democratic People’s Republic of Korea” was replaced with “North Korea”).

The confidence reporting scale was entered using the 1-5 keys on the computer keyboard as shown in Figure 2.8. As with the click train task, an “error” response was available if subjects had such low confidence they believed they had erred. The outcomes on these responses were not statistically differentiable from chance accuracy using a binomial 95% confidence interval. Since subjects could not predict their errors, perceived error responses were counted as lowest confidence responses. Of 1600 trials, only 34 were “error” responses.

To encourage participation by reducing the clerical demand on subjects, I programmed the quiz to automatically Email the results back to a dedicated Email account for the study (Gmail, Google). The software checked for an Internet connection on launch, and if not detected, informed the subject that Internet access was required to participate.

On the first run of the program on a new computer, the software created a unique client ID file that contained a randomly generated ID number. All subsequent sessions from the same computer would be marked with the ID code. Subjects were requested not to share their computer, and were given instructions explaining how to delete the client ID file if sharing was necessary.

Procedure

Subjects viewed an animated title screen introducing the quiz while it loaded. Subjects were then required to register for the study by entering their age, gender, handedness and indicating that they had read the informed consent document. Subjects then read the instructions (figure 2.8) and clicked the arrow to begin the session.

Each session was 100 trials in length. No subject completed more than one session. Individual countries for comparisons were selected randomly without replacement from the list of 51 countries, while ensuring that no comparison was shown twice within a session. No feedback was given about whether choices were correct or incorrect, to maintain that the quantities being compared were general knowledge and not related to recent learning within the session. The time course of trial events is shown in figure 2.9.

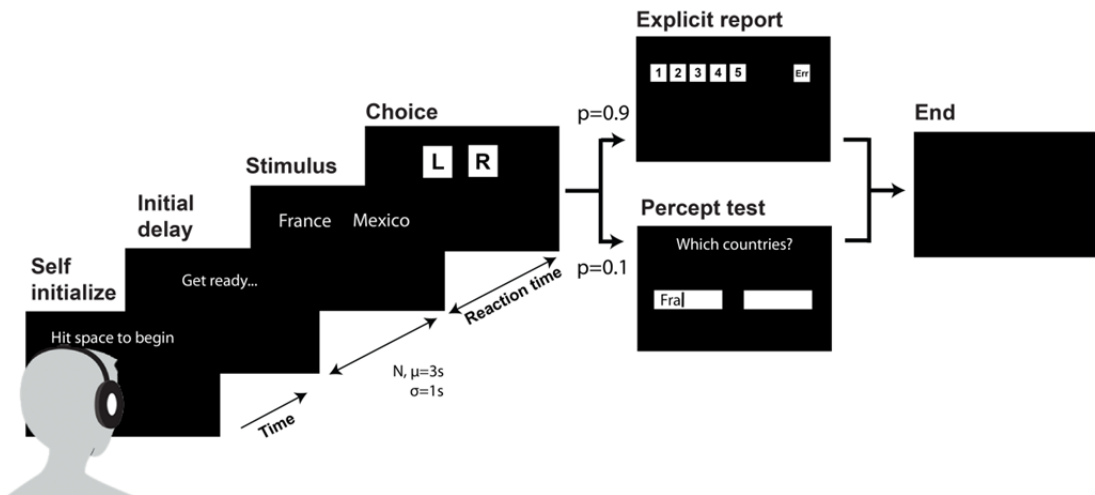


Figure 2.9: Flow of trial events for Confography task. Subjects initialized trials by pressing the space bar. After a random delay (Gaussian, mean = 3s, var = 1s) two countries were displayed. The subject had three seconds to indicate the larger country using the “o” and “p” keys to indicate left and right respectively. Subjects then entered either entered a confidence report, or recapitulated the names of the countries they had just compared.

To verify that subjects had perceived the correct countries, 10% of trials were probe trials where subjects were asked to type the names of the countries they had just compared in lieu of providing a confidence report. The probe trials were randomly chosen with the constraint that no two sequential trials could be probe trials. If subjects did not respond within three seconds, the screen displayed “Too slow” accompanied by a negative feedback sound. Upon completing the game, subjects were informed that their data was being transmitted. Subjects were then notified if the transmission was successful, and the quiz automatically closed. We provide the results of the Confography study in section 3.4.

2.1.5 Design of a real-time click generator and response capture device

A wide range of applications in Neuroscience require patterns of precisely timed electrical pulses. Some of these are classical electrical stimulation, control of lasers in optogenetics, control of sensory stimuli in psychophysics and synchronization of ensembles of instruments in electrophysiology. Often, these pulse patterns must be initiated with high precision by a signal from a governing instrument. In some cases, this can be accomplished with high end computer I/O cards (i.e.

National Instruments), commercial stimulators (i.e. Master 8, AMPI) or arbitrary waveform generators (i.e. 53200A, Aligent) – but each of these solutions can cost thousands of dollars, posing a challenge for high throughput assays in settings with limited funding. More importantly, the latter two options are difficult to interface with MATLAB based experimental control systems.

To provide a better option for these applications, I developed Pulse Pal, a programmable pulse and waveform generator. Pulse Pal can be assembled at a soldering bench in under an hour for a cost of under \$250 in common electronic parts available from major industry suppliers (figure 2.10a). The Pulse Pal design files and software are open source, and publically available at <https://bitbucket.org/LucidBiosystems/pulsepal>. The device was intended to be part of a larger open source bioscience instrumentation project I termed Lucid Biosystems in web based documentation.

Pulse Pal delivers trains of monophasic or polyphasic square pulses ranging from -10V to 10V with 8 bits (256 divisions) of voltage resolution. Pulses as short as 50 μ s can be reliably delivered with less than 2 μ s of jitter. Pulse Pal has two optically isolated digital trigger channels and four analog output channels. Trigger channels sense TTL triggers, and can be arbitrarily mapped to control specific output channels in software. Each output channel can be independently programmed with a different pulse train, and each trigger channel can be programmed to respond differently to logic.

Pulse Pal communicates directly with a client in MATLAB, which allows easy programming and software triggering. Using the client, researchers can easily create parametric pulse trains by specifying pulse width, amplitude, inter-pulse interval, playback delay and train duration, and several more advanced parameters for common physiology assays. In cases where parametric pulse trains are constraining, it is also possible to specify the onset time and voltage of each pulse in two arbitrary trains of 1,000 pulses, which can be loaded for playback on any output channel. Arbitrary pulse trains can be looped to create continuous sine waves, white noise or other simple acoustic stimuli for behavioral feedback. Using an EEPROM memory chip, Pulse Pal is programmed to retain programming between power cycles. It can be operated as a stand-alone device, drawing all of its power from an AC USB accessory charger or a USB battery. In this mode, Pulse Pal is programmed

by a clickable mini-joystick for navigating menus and changing parameters displayed on an OLED screen (Figure 2.10b). Its enclosure is cut from a single acrylic sheet on a commercial laser cutter, and assembled with screws, nuts and stand-offs. The enclosure can be ordered online from most laser cutting services.

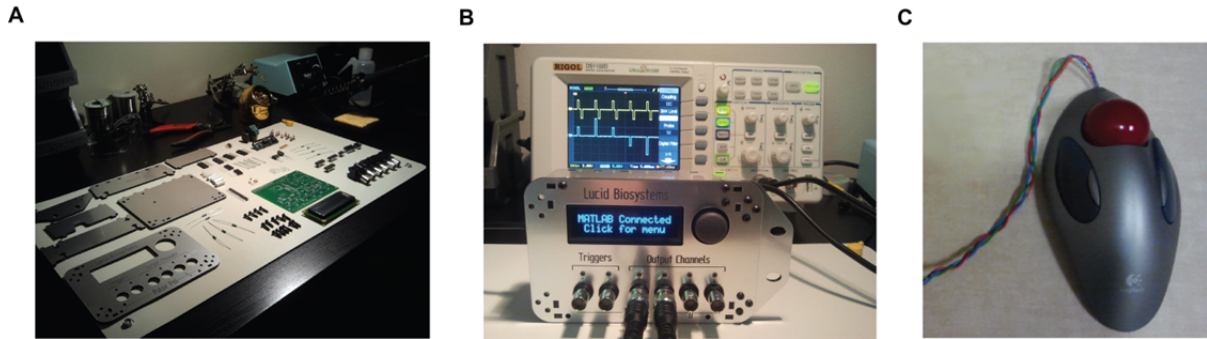


Figure 2.10: The Pulse Pal device and custom interface for human responses. **A:** Pulse Pal can be assembled at a soldering bench in one hour, using common parts. **B:** A bench test of Pulse Pal, showing different trains of 1ms unipolar and bipolar pulses simultaneously delivered on each of two channels. **C:** The response device for the human experiments was a modified computer mouse. Wires were run from Pulse Pal’s 5VDC supply to the proximal contacts on the two lateral buttons, and from the distal button contacts back to Pulse Pal’s trigger channels. Subjects used the buttons to indicate left and right responses. The red trackball component was unused.

A detailed account of Pulse Pal’s circuit design is available in the repository. In brief, Pulse Pal is controlled by an open source ARM Cortex M3 microcontroller platform (Maple, Leaf Labs). The microcontroller interfaces with the Pulse Pal MATLAB client via a virtual serial port, and writes voltage instructions in serial byte streams to a bipolar digital to analog converter IC (MAX500, Maxim) to control the four output channels. To power the DAC, An integrated boosting voltage regulator module (PT5061A, Texas Instruments) provides a bipolar 12VDC power supply from the unipolar USB 5V supply.

Pulse Pal was used in our human and rodent decision making experiments to generate precisely timed trains of Poisson clicks. Prior to each trial, the onset times and durations of each click were uploaded to the two memory slots for custom pulse trains. Pulses were specified to be bipolar, with each phase lasting 100 μ s, and amplitude was calibrated to equalize click volume for the speakers

in each apparatus. Trigger channels were set to “toggle mode”, which interpreted incoming pulses during stimulation as stop signals. When triggered by the behavior system via TTL pulse, Pulse Pal delivered sound directly to speakers until the rat withdrew, causing the behavior system to send the second TTL stop signal.

For the human study (section 2.1.2), no real-time behavior system was used. A custom modification to the default Pulse Pal firmware allowed it to fulfill this role. Pulse Pal was outfitted with a two button response device (figure 2.10c) and programmed with three additional functions: 1. to receive the stimulus delay, reward delay and correct response from the governing computer, 2. to terminate the click train on button press and enter a “time investment” mode which was terminated by a timer if the subject chose correctly, or otherwise by release of the response button, and 3. to return the choice, reaction time and time investment to the governing computer after each trial. Since Pulse Pal’s main loop cycles exactly every $50\mu\text{s}$, the same resolution applies to these measurements with respect to the click train heard by the subject – a standard that is virtually impossible with purely software based tools.

At the time of this report, 21 Pulse Pal devices have been adopted to empower 9 research projects in 3 different research groups. In some of its present roles in ongoing research, commercial alternatives at the scale deployed would have been cost prohibitive, demonstrating the capacity of open source hardware like Pulse Pal to change which possibilities researchers think to entertain. I anticipate that with time and associated publications, Pulse Pal and related open source instrumentation projects (such as Open Ephys; Josh Siegle and Jakob Voigts, MIT) will become transformative technologies in our field.

2.2 Data analysis

Our decision confidence experiments in humans and rats were designed with the intent to relate measures between species. Raw measures describing each trial were: a stimulus onset delay, two vectors of click times, a reaction time, a time investment, a choice outcome and indication

whether reward was delivered. Specialized data in the human task variants included explicit confidence reports and explicit confidence response time. Human data lacked click times in the case of the Confography task (section 2.1.4) and instead had country populations as evidence.

Statistical analysis of our data had two overarching goals: 1. to characterize the relationship of decision confidence to raw and derived measures of behavior, and 2. to evaluate the fitness of a computational model that can explain these relationships in our data. Establishing relationships between variables (for instance, confidence and reaction time) generally relied on standard statistical tests for mean differences and linear regression (section 2.2.1). Determining model fitness required methods described in section 2.2.2.

All data were analyzed with MATLAB R2011a (Mathworks) using the MATLAB statistics toolbox and custom analysis code.

2.2.1 Statistics

Statistical tests often make assumptions about the shape of the distribution from which a sample was drawn, in order to exploit properties of that distribution. In virtually every analysis in the present work, I was fortunate to have sufficient computational power to use bootstrap methods in lieu of parametric tests or confidence interval estimates – avoiding the caveat of unknown distribution shape, and vastly simplifying the complexity of some analyses. In this section, I briefly describe how each bootstrap statistic was computed.

Bootstrap confidence intervals

We frequently needed to determine our confidence that a true parameter θ is within a fixed interval of a sample parameter estimate $\hat{\theta}$. We first describe how this is accomplished for the error of means, and then generalize to more complex statistics.

We computed bootstrap confidence intervals for means using the Monte Carlo method described in (Efron and Tibshirani 1993). A brief description of this method is as follows: 1. from our

sample distribution of measurements, we drew measurements independently and with replacement to create a new sample distribution of measurements the same size as our original sample. We repeated this process $B = 1,000$ times (unless otherwise indicated) to create a distribution of 1,000 bootstrap samples, X^* , as an estimate of the *empirical* distribution of samples \hat{P} that would arise as $B \rightarrow \infty$:

$$X_1^{*(b)}, \dots, X_n^{*(b)} \sim \hat{P} \quad b = 1, \dots, B \quad (2.2)$$

2. We computed the mean of each sample, to generate a distribution of replicate means.

$$\hat{\theta}^{*(b)} = \text{mean}(X^{*(b)}) \quad b = 1, \dots, B \quad (2.3)$$

where $\hat{\theta}^*$ is the distribution of replicate means.

3. To compute confidence intervals from this distribution, we computed the interval of $\hat{\theta}^*$ containing the central 95% (unless otherwise indicated) of the density of means.

Often, we needed to determine confidence in a statistic computed from multiple paired measurements (for instance, averaged across all sessions, the slope of the best linear regression fit of reaction time vs. accuracy, or the significance of AUC in ROC analysis). To compute confidence in these cases, we drew trials at random with replacement to create B bootstrap samples of trials, each the same size as our entire dataset. For each bootstrap sample dataset, we computed the statistic of interest, and iterated to create a replicate distribution of B replicate statistics, determining confidence from the interval containing 95% of the replicate statistic density.

Bootstrap test for difference between means

A test for the difference between means of two samples yields a p-value describing the probability that the difference in means of the two underlying distributions is not equal to 0. Unlike the case of statistics computed from paired measurements, computing the significance of the difference between a parameter in two conditions requires bootstrap samples, $X1^{*(b)}$ and $X2^{*(b)}$ to be drawn independently from each condition's original sample. The size of each bootstrap sample was

equal to the total number of measurements for each condition. After iterating to create B total samples (X^*), we computed the difference between means of each pair of bootstrap samples to create a distribution of replicate differences:

$$\hat{\theta}^{*(b)} = \text{mean}(X1^{*(b)}) - \text{mean}(X2^{*(b)}) \quad b = 1, \dots, B \quad (2.4)$$

The test passed if an interval containing 95% of the density of replicate mean differences did not contain 0. To compute the p-value of the test, we subtracted the mean of $\hat{\theta}^*$ and computed the absolute value of the resulting distribution, to create a distribution of differences between replicate mean differences and their average. The fraction of this distribution greater than $\text{abs}(0 - \text{mean}(\hat{\theta}^*))$ is the p value of the two-tailed, two-sample test for difference between means[].

2.2.2 Model fitting and measuring goodness of fit

Fitting psychometric data

Psychometric data relating discriminability to choice accuracy are typically fit with Weibull functions using maximum likelihood, to determine classification performance and choice bias (Wichmann and Hill 2001). For modeling a two-choice decision (chance accuracy = 50%), a Weibull function can be expressed in terms of two free parameters defining the threshold for 80% performance and slope (Fine 2009):

$$y = 0.5 * e^{-\left(\frac{kx}{t}\right)^b}; \quad k = -\log\left(\frac{1-a}{0.5}\right)^{\frac{1}{b}} \quad (2.5)$$

where a is the performance level that defines threshold (0.8), and free parameters t and b define the threshold and slope of the function. For a given set of parameters, a likelihood function was evaluated with respect to data to determine the likelihood of the fit for subsequent maximization. Likelihood was computed as log value using the following equation:

$$L = \sum r_i \log(W_i) + (1 - r_i) \log(1 - W_i) \quad (2.6)$$

Where W_i is the value of the Weibull function evaluated at the stimulus intensity of trial i , and r_i is the trial's outcome (1 if correct, 0 if incorrect). Thus, if a correct trial co-occurred with a high probability predicted by the Weibull function, or an incorrect trial was co-occurred with a low probability prediction, a large value was added to L . To determine the best fit parameters, the negative Likelihood was minimized using the Nelder-Mead simplex algorithm (Lagarias, Reeds et al. 1998) (default for MATLAB r2011a function `fminsearch`).

In cases where a bipolar psychometric function was used (probability of classification vs. balance of intended evidence), curves were fit with probit regression using MATLAB function `glmfit`.

2.2.3 Reverse correlation analysis

Reverse correlation is a method used in neurophysiology to determine the receptive fields of neurons, by computing the average stimulus that preceded an action potential (Aertsen and Johannesma 1981; Ringach and Shapley 2004). This method was subsequently adopted in decision making research to determine which features in spatially or temporally structured evidence contributed to choice (Neri, Parker et al. 1999; Nienborg and Cumming 2009; Zylberberg, Barttfeld et al. 2012). To compute the reverse correlation of Poisson click stream evidence, chosen and anti-chosen click trains were separately aligned to the moment of choice on evidence-neutral trials (50Hz/50Hz). On each trial, these evidence streams were partitioned into 100ms bins, the number of clicks in each bin was counted, and the baseline click rate (5 clicks/100ms bin) was subtracted from the count to compute excess clicks. On each trial, one bin was partially visited since reaction time did not fall precisely on the edge of a bin boundary. For this bin, a corrected baseline was computed as 50Hz multiplied by the fraction of the bin visited. Once excess clicks in each bin were computed for each trial, we computed the mean and 95% confidence interval of excess clicks in each bin. The resulting time courses for chosen and anti-chosen evidence were plotted, showing epochs with statistically significant enrichment or paucity of clicks (Section 5.4). To determine how evidence contributed to confidence, separate reverse correlations were computed and superimposed for high

and low confidence trials, showing epochs where evidence differed between these conditions (Section 5.5).

2.2.4 ROC analysis to determine how well evidence predicts behavior

Receiver Operator Characteristic (ROC) analysis is a method used in signal detection theory to determine how well a binary classifier makes predictions (Peterson, Birdsall et al. 1954; Green and Swets 1966). For a set of trials belonging to either categories A or B with a predictive characteristic C for each trial in a continuous range, the classifier determines which category each trial belongs to by comparing each trial's C value to a threshold. As the threshold is varied across the range of inputs, the *ROC curve* is constructed by plotting the fraction of correctly classified group A trials (true positive, ordinate), against the fraction of group B trials which were mislabeled group A (false positive, abscissa). The area under the curve (AUC) then provides a measure of the classifier's performance, where 0.5 = chance discrimination accuracy, and 1 = perfect discrimination. We used ROC analysis to determine how well the balance of left and right clicks predicted left and right choice, and how well the balance of chosen-side and anti-chosen side clicks predicted confidence in different time windows. Time preceding choice was divided into windows, typically 100ms in size. For each window, the ROC curve was constructed and AUC estimated from the curve by trapezoidal approximation (using the MATLAB function `trapz`). To determine the confidence interval of the AUC statistic, bootstrap replicates of the AUC were calculated for 1,000 datasets the same size as our original, consisting of trials (click train-confidence pairs) drawn with replacement (see section 2.2.1). Time windows with AUC values significantly larger than 0.5 indicated that click balance during that time predicted choice (or confidence) with greater than chance accuracy.

Chapter 3

Verbal confidence recapitulates the brain's statistical confidence in its assertions

The brain is fundamentally an information processing engine, which must constantly approximate the true state of the world using samples of noisy sense data. In this, the brain's challenge is fundamentally statistical, and can be solved in the abstract using Bayesian inference – a process that generates a measure of confidence in a prior classification. We wondered whether the human feeling of confidence provides conscious access to a mental variable that is equivalent to confidence in statistical inference. To test this, we evaluated human subjects in a pair of tasks that measure confidence in a subject's classification of quantities in different modalities. From the normative definition of confidence as a conditional probability, we derived predictions describing the interrelationships between confidence, evidence, and classification accuracy. Among these predictions is a counterintuitive property of confidence – that the lowest average confidence occurs when classifiers err in the presence of the strongest evidence. We show that these predictions are true of p-values generated by parametric and non-parametric statistical tests, using a range of noise patterns to corrupt primary evidence. We also report that these same predictions characterize human verbal confidence reports in two markedly different decision making tasks, revealing a novel and counterintuitive property of confidence in humans.

3.1. Normative confidence predicts interrelationships of confidence, evidence and accuracy

The title of the present thesis is “A framework for understanding decision confidence”. The premise of our framework is a normative model of confidence which relates confidence to evidence with a conditional probability. We take as first principles that confidence is a probability estimate describing a belief, and that confidence is related to the available evidence supporting the same belief by a conditional probability (Kahneman and Tversky 1972). As such, Bayes rule provides a way to

understand confidence in terms of quantifiable evidence (Ferrell and McGoey 1980; Griffin and Tversky 1992).

$$C = p(D|E) = \frac{p(E|D) * p(D)}{p(E)} \quad (3.1)$$

In terms of *decision* confidence, C is confidence expressed as a probability estimate that the chosen option is correct, D is the hypothesis that the chosen option is correct, and E is the perceptual evidence used to decide.

From this normative definition of confidence, equations describing the predicted interrelationships between confidence, choice accuracy and evidence were formally derived by Balazs Hangya, a postdoctoral fellow in our research group. These derivations, provided in Appendix I as supplementary material, make several predictions about confidence describing choices among options. In brief, these equations show that when confidence describes a set of two-choice decisions made with a continuous range of evidence strengths that does not vary in time, the following predictions hold: 1. Confidence is directly correlated with choice accuracy. In instances where the confidence signal has not been scaled or transformed, this relationship is an exact identity. 2. Confidence is correlated with *discriminability* for correct classifications, and anti-correlated with discriminability for incorrect classifications, where discriminability is defined for a given stimulus and category boundary as an ideal observer's probability of classifying the stimulus correctly. 3. Even at a fixed discriminability, confidence predicts choice accuracy. To visualize these properties of confidence, a simulation using the equations derived in Appendix I was further contributed by Balazs Hangya. The patterns produced by the simulation are shown in Figure 3.1.

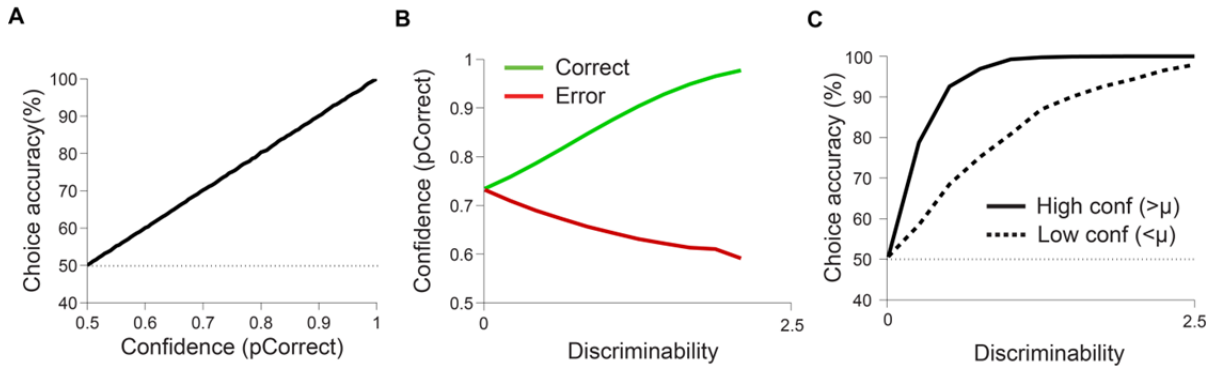


Figure 3.1: Three patterns in decision confidence derived from the normative model. Individual simulations of the equations in appendix 1 produced a profile of three patterns in confidence. **A:** Confidence predicts accuracy. Here, confidence is shown as a probability estimate that a choice is correct. At 0.5, confidence is lowest, and predicts chance accuracy. At 1, confidence is highest and predicts perfect accuracy. **B:** Across the range of the strength of evidence provided to a decision maker, confidence is correlated with discriminability for correct classifications, and anti-correlated with discriminability for incorrect classifications. Evidence strength for each hypothesis on individual decisions was modeled as a value drawn from a Gaussian distribution with variance 0.6, and mean differences (*Discriminability* on figure axes) ranging from 0 to 2.5. On trials where the weakest evidence is delivered (0), mean confidence is counterintuitively *half* of its maximum value. Also counterintuitively, the *lowest* confidence is predicted to occur on errors in the presence of the *strongest* provided evidence. **C:** When a psychometric function relating evidence strength to accuracy is split by the high and low halves of the range of confidence values, confidence still predicts accuracy for a given evidence strength – high confidence predicts high accuracy and low confidence predicts low accuracy. When evidence is very strong, no errors are made and the curves converge. When evidence is neutral, no information is available to inform choice regardless of confidence, and the curves also converge.

Two aspects of the second pattern are counterintuitive. Firstly, if confidence and accuracy are an identity as per the first derived pattern, when provided with neutral evidence that results in chance accuracy, intuition suggests that confidence should be *lowest* – not averaging half its range. Secondly, when provided with the strongest evidence in a given range, intuition suggests that confidence will be highest regardless of the choice outcome (which is not yet available when confidence is computed).

The key to understanding these patterns is that in a confidence computation, there are *two* measures of evidence strength – 1. the strength of *provided evidence* supplied to a decision maker, and 2. The strength of *perceptual evidence*, the evidence apparent in the decision maker’s frame of reference, which occurs when an imperfect copy of the provided evidence is acquired by sampling.

Both intuitions are true for *perceptual* evidence – when the strength of perceptual evidence used by a discriminator is weakest, confidence is lowest and when strongest, confidence is highest.

To understand the first of these counterintuitive properties, that confidence is not lowest when *provided evidence* is lowest, it may be helpful to consider the case of a decision maker provided with perfectly neutral evidence for each decision in a series. While the provided evidence is precisely neutral, the decision maker only has access to *samples* of the perfectly neutral evidence, which individually have a value different from neutral. The decision maker learns that any experienced strength in the resulting distribution of perceptual evidence strengths predicts chance accuracy, and thus he has the lowest possible confidence for all strengths resulting from neutral evidence. In a different instance, the decision maker is only provided with evidence of a precisely fixed positive strength. The decision maker learns that the distribution of experienced evidence from *this* source predicts accuracy 65% of the time. When provided evidence strengths are delivered at random from a continuum that ranges from neutral evidence to evidence that is classified nearly perfectly, the perceived strength of evidence experienced on a trial when neutral evidence is *provided* overlaps with the distributions of perceptual evidence strength produced by informative stimuli. Thus the appropriate reward probability estimate for perceptual evidence strengths generated by neutral stimuli is greater than chance, and increases linearly with delivered evidence strength to an asymptote at $p=1$ as the decision maker stops making errors.

The second counterintuitive property, that confidence is lowest in the presence of strong evidence when a decision maker errs, results from dividing the distribution of perceptual evidence strengths into separate groups for correct and incorrect classifications. In the subset of perceptual evidence strengths generated on trials with very strong provided evidence, the perceptual evidence will support the wrong hypothesis with a very low probability. The average amount by which perceptual evidence supports the wrong hypothesis when it does so, declines with increasing strength of provided evidence as the probability of these rare events also declines. Thus an error on a trial with

very strong delivered evidence produces weak perceptual evidence almost always, while a neutral evidence error produces a higher *average* strength of perceptual evidence.

According to the derivation, the three patterns in Figure 3.1 are initially true of all realizations of decision confidence. However, actual confidence signals captured from environmental sources can be expected to display various transforms of these patterns. For instance, a continuous confidence signal mapped onto a discrete reporting scale may exhibit apparent compression along the confidence axis in the first two projections of Figure 3.1, if the mapping function does not map confidence to the full range of the scale. Furthermore, if a sufficient amount of noise is added to a confidence signal, it will be undetectable. To confirm the generality of the predicted patterns in *uncorrupted* confidence signals, we chose to determine whether these patterns apply to common statistical confidence values: p-values of two-sample significance tests.

3.2. Statistical confidence measures are characterized by normative model predictions

In statistical hypothesis testing, p-values are fundamentally confidence values which are inverted to describe uncertainty by convention; estimates of the likelihood that a null hypothesis is correct. We reasoned that if the patterns we derived from the normative model are general properties of confidence, they should also describe the behavior of $(1 - p)$ -values. Two common tests provide p-values which: 1. estimate the likelihood that a sample was drawn from a distribution whose mean is either larger than a fixed value (one-sided one sample t-test) or 2. estimate the likelihood that a sample was drawn from a distribution whose mean is larger than the mean of an alternative sample (one-sided two-sample t-test, one-sided bootstrap test for difference between means (Manly 2007)). The t-tests are parametric tests and require that the populations and samples be normally distributed, while the bootstrap test is non-parametric and makes no assumptions about the sample and population distributions.

I created simulations to determine whether these tests exhibit the patterns predicted from the normative model. For the one sample t-test, on each of 10 million trials, a sample of 30 measurements was drawn from a distribution. On each trial, the mean of the distribution was offset from 0 by a random amount drawn from a uniform distribution between -0.325 and 0.325. The absolute difference between the population mean and 0 was defined as discriminability. Trials were scored as correct if both the population mean and sample mean had the same ordinal relationship with respect to 0. On each trial, a one-tailed single sample t-test was performed to measure the probability that the ordinal relationship informing choice was correct (that population mean had the same ordinal relationship with respect to zero as the sample mean). The p-value produced by this test is shown with respect to discriminability and outcome in Figure 3.2a-c, showing unambiguously that this statistical confidence measure displays the predicted interrelationships with the choice outcome and discriminability measures.

For the two-sample t-test, on each of 10 million trials, two samples of 30 measurements were drawn from two normal distributions of variance 1 (A and B). The means of distributions A and B were offset from 0 by random amounts drawn from a uniform distribution between -0.5 and 0.5. For each trial, the absolute difference between these offsets was a measure of discriminability. On each trial, if the mean of sample A was larger than the mean of sample B, we tested the hypothesis that the mean of distribution A is larger than the mean of distribution B and vice versa. I scored each trial as correct if this was true of the population based on our offsets, and incorrect if false. On each trial, a p-value was computed using one-tailed variants of each of the two-sample hypothesis tests listed above, providing confidence that the ordinal relationship between sample means was also true of the population means. The results of this simulation are shown in Figure 3.2d-f.

Since the bootstrap test does not depend on an assumption that evidence is normally distributed, I used underlying distributions A and B which were *exponentially* distributed. Offsets for the population means were uniform, ranging between 0 and 1, and the bootstrap sample size was

1,000. Due to computational constraints, only 2 million trials were simulated. The resulting patterns relating bootstrap p-value to accuracy and discriminability are shown in Figure 3.2g-i.

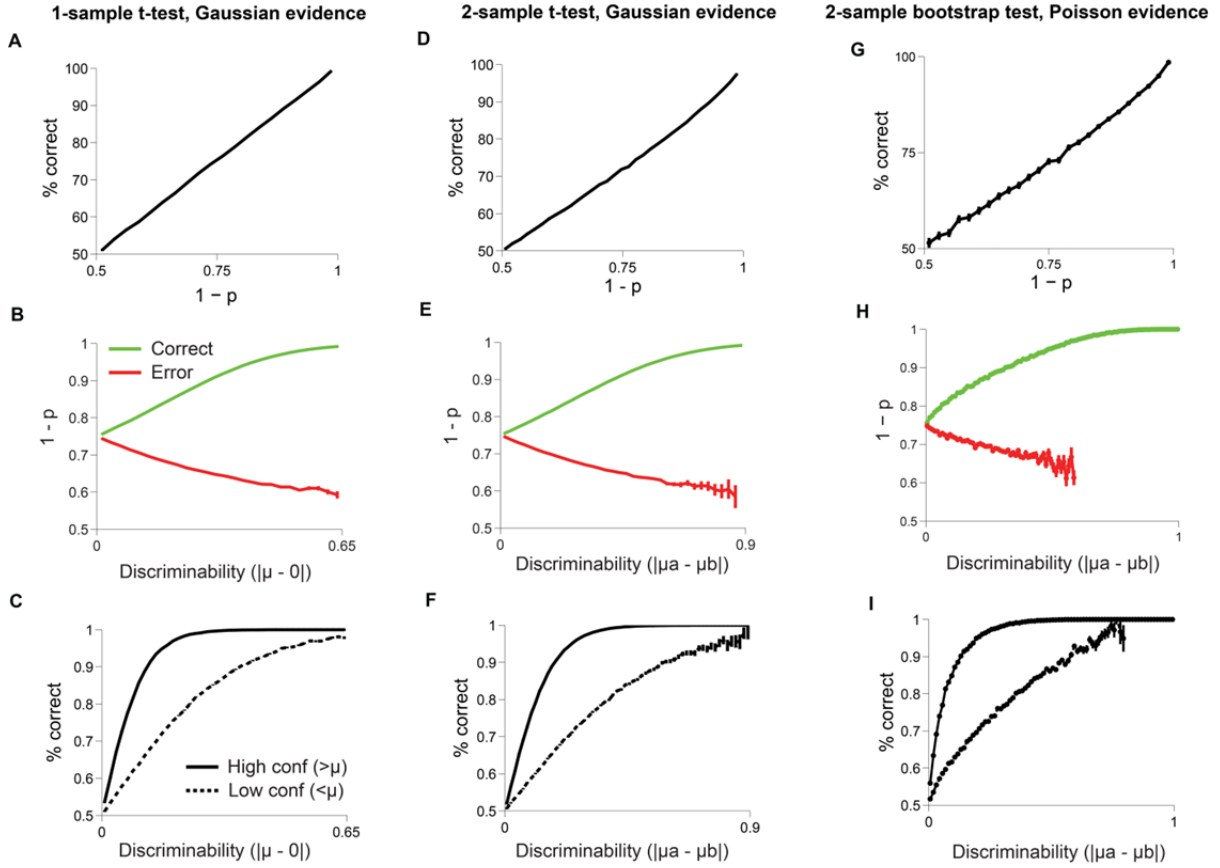


Figure 3.2: Statistical confidence measures reproduce patterns predicted by the normative model. **A-C:** A simulation of 10 million trials evaluated the one-sided, one sample Student's t-test p-value with respect to accuracy and discriminability. Discriminability was defined as the absolute difference between the population mean and zero. Population distributions were Gaussian, with means ranging from -0.325 to 0.325. Trials were scored as correct if both the population and sample mean had the same ordinal relationship with respect to 0. Since estimated probability p is an uncertainty measure by convention, it is displayed as $1-p$ to indicate confidence. In this form, the p -value displays an identity relationship with accuracy and is both correlated with correct-trial discriminability, and anti-correlated with error-trial discriminability. A pair of psychometric functions was evaluated for trials with greater and less than average confidence values (solid and dotted lines respectively in panel C) showing that p -values contain predictive information about outcome even at fixed discriminability. **D-F:** The two-sample Student's t-test p -value was evaluated with Gaussian evidence as for the one sample analysis with two changes: discriminability was defined as the absolute difference between the two population means, and outcome was scored correct if the sample means preserved the ordinal relationship between population means. **G-I:** A similar simulation evaluated the p -value in a one-sided bootstrap test for an ordinal relationship between means. The p -

value is shown with respect to accuracy and discriminability. Population distributions were exponential, with means ranging from 0 to 1. As in the parametric example, the p-value is correlated with accuracy and correct-trial discriminability, and anti-correlated with error-trial discriminability. Plot E is shown at higher resolution than plot B to demonstrate that correct and error p-values converge as discriminability approaches 0. As with the t-test, bootstrap test p-values contain information about outcome even at fixed discriminability (panel I).

3.3. Human verbal confidence in percept classifications approximates statistical confidence

We found our set of three derived patterns to be an excellent descriptor of trends in both normative Bayesian confidence and Frequentist p-values. The latter two tests shown in Figure 3.2 explicitly estimated the likelihood that an inferred ordinal relationship between two samples is true of their underlying populations. In two-choice perceptual decision making, the decision maker is tasked with a similar statistical problem – to gather noisy samples of perceptual evidence supporting each alternative, and using the greater measurement, infer the ordinal relationship between their environmental sources. We wondered whether the patterns of statistical confidence would be evident in the *feeling* subjects individually identify as “confidence”, when asked to provide this report after making a choice among two options. To test this, we designed two markedly different decision making tasks – one basing choices on two independent streams of precisely controlled sensory evidence, and the other requiring subjects to gather evidence from memory, to compare two quantities that are common knowledge. We present our findings from the sensory task in this section.

To determine whether human verbal confidence reports also reproduce normative model predictions, we tested four human subjects, designated Subject H1 – Subject H4, in the explicit-only variant of our confidence task (section 2.1.2). On each trial, following a random delay, subjects listened to a pair of Poisson click streams, one stream delivered to each ear, and indicated which stream was clicking faster on average by pressing one of two horizontally oriented buttons. The click streams were terminated the instant the choice button was pressed. Subjects could sample up to three seconds of clicks before deciding, and forfeited the trial’s reward if they responded after this deadline. Immediately after pressing the choice button, subjects were prompted to report their feeling of

confidence in their choice on the 1-5 scale by pressing a button on a keypad (Figure 2.2). Following the report, subjects were visually informed if they were correct, and had to initialize the next trial by pressing a button.

To measure decision confidence based on evidence of varying strength, we generated Poisson click stream pairs with 100 clicks per second summed between left and right channels, and a range of click rate balances from 50Hz/50Hz (zero discriminability) to a typical balance of ~65Hz/45Hz at which the subject chose the faster click train with 90% accuracy.

In this section, we first determine a metric of the strength of evidence actually delivered to subjects on individual trials, as a more meaningful measure of the discriminability of evidence contributing to individual confidence reports than intended Poisson rate (section 3.3.1). We then determine how much of the final moments of a click stream was ignored because subjects had committed to a button press response, to provide a more relevant window for computing our discriminability metric, and estimate what fraction of the post-commitment window was actually required for the button press motor response (section 3.3.2). To determine how accurately and reliably subjects performed, and to what extent biases contributed to individual decisions apart from sensory evidence, we measured psychometric performance in section 3.3.3 and chronometric performance in section 3.3.4. In this precisely characterized behavior, we show the robust presence of statistical confidence patterns in human confidence reports, and characterize these reports further in the remainder of section 3.3.

3.3.1 Discriminability measures with free-response Poisson click evidence

To characterize the relationship of discriminability to confidence, reaction time and choice, we sought to determine the best metric of Poisson click stream evidence to relate these measures on individual trials. For our perceptual confidence tasks, we implemented graded evidence strength in a manner similar to the olfaction based decision making studies of Kepecs et al 2008 and Felsen et al 2008, by creating several discriminability categories (neutral-evidence, hard, medium and easy). We

adjusted the evidence strength for each subject, until performance in each of the 3 non-neutral categories was as evenly spaced as possible between 50 and 100% choice accuracy. For studies with static stimuli, this method permits a researcher to analyze the relationship of discriminability to other variables in a limited size dataset without needing to create discriminability bins of arbitrary size. However, the use of discrete categories posed a problem for our precise, temporally structured stimulus as illustrated by the example stimulus in figure 3.3.

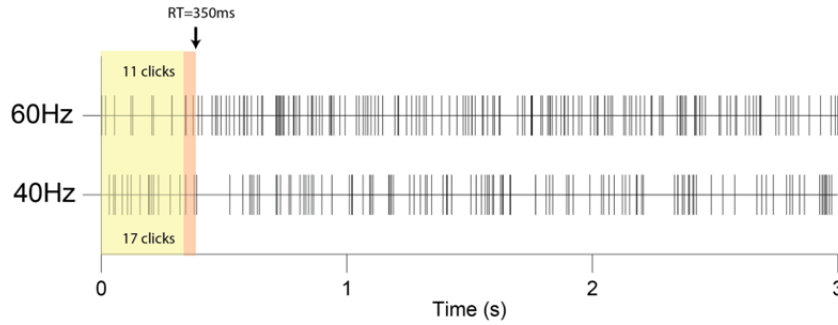


Figure 3.3: In a free response task, underlying rate may not capture the strength of Poisson evidence gathered on individual trials. Two three-second Poisson click trains are shown in an example trial where the subject responded after 350ms. In this 350ms window (combined yellow and red shade), *provided* evidence strongly supports the incorrect decision; click count for the 40Hz Poisson train is over 50% larger than for the 60Hz train. Additionally, the subject may not have made use of the final moments of evidence (shaded in red, supporting evidence discussed below), due to the time needed to execute a motor response following choice. However, evidence collected during this “post-decisional” period could have contributed to subsequent confidence reports.

The example in figure 3.3 is atypical, since such a large click count difference in the opposite direction of the intended rates ought to occur with low probability. We sought to determine the extent to which underlying rate was a useful discriminability category boundary in our dataset. A metric of the balance of sampled evidence, β , was computed for an ideal observer with equation 3.2,

$$\beta = \frac{nL - nR}{nL + nR} \quad (3.2)$$

where nL is the number of left clicks in the evidence stream, and nR is the number of right clicks. If all clicks occur on the left side, $\beta = 1$ (or on the right side, $\beta = -1$); and $\beta = 0$ if left and right click counts are equal. This measure is useful for evaluating both experienced evidence, and also for

evaluating *intended evidence* since the 3-second click streams computed were truly Poisson (section 2.1.1). Thus unlike in Kepecs et al 2008, where odor mixtures are delivered at precise ratios to the sensory periphery on each trial, even our full length stimuli have experimentally designed variability.

If the evidence experienced on a trial by trial basis *perfectly* reflected the intended difficulty categories on individual trials, the distribution of the absolute value of β (excluding neutral-evidence trials) would consist of only three values. With a small amount of variability on individual trials, the expected distribution would be tri-modal. With a sufficiently large amount of noise to obscure the category boundary definitions, the distribution would appear unimodal. This measure is shown for the entire 3 second evidence streams pre-computed (Figure 3.4a), and separately for the portion of those same streams actually experienced (Figure 3.4b).

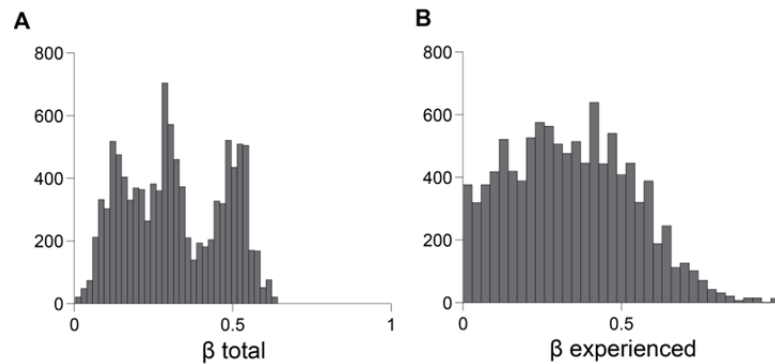


Figure 3.4: The balance of experienced evidence forms a continuum despite intended difficulty categories. For a single example subject, evidence balance metric β is shown for the 3 second Poisson click trains generated with intended easy, medium and hard evidence strengths (A) and for the portions of the same trains actually experienced (B).

We sought to determine a post-hoc measure of discriminability that reflected the strength of evidence used by the subject. Since experienced evidence does not reflect intended categories on individual trials, we sought to exploit the ability of a binned measure to capture the range of discriminability at higher resolution. In figure 3.5, I show psychometric functions for two example subjects using each of these measures, showing how using β as an evidence measure allows us to study trials with intermediate levels of discriminability.

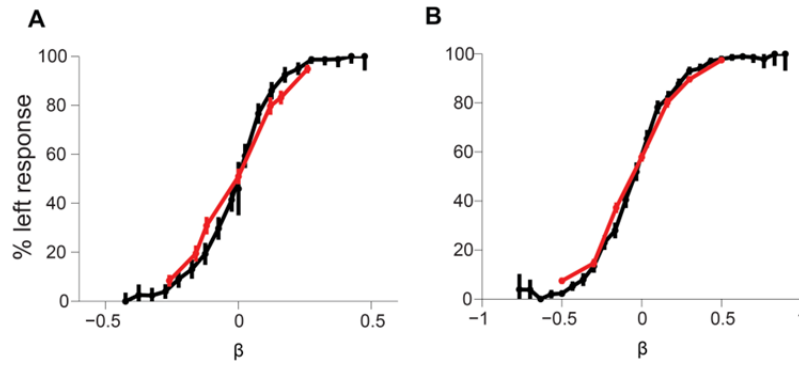


Figure 3.5: A measure of post-hoc discriminability, β captures the full range of graded performance. Data from two human test subjects (panels A and B) are shown. The red line shows a 7-point psychometric function using underlying rates as evidence. The value of β was calculated for underlying rates using equation 2.1, with left and right rates replacing n_L and n_R respectively. The black line shows β calculated using experienced evidence (as in equation 3.2). Error bars show 95% binomial confidence interval.

3.3.2 Estimating the time of motor response commitment and motor response latency

Since our task is a free response task, the button press or snout withdrawal motor responses we capture can be assumed to occur shortly after the instant of a subject's commitment to a choice. Thus, decisions may be insensitive to evidence in the final moments of the ongoing stimulus, which is terminated by the motor response. To measure the evidence that informed individual decisions in our task, we wanted to compute a derivative measure of β using only the click balance experienced by the subject, and disregarding the clicks that did not contribute to choice. We refer to this measure as β' . For this reason, we sought to approximate which portion of the final moments of the stimulus did not contribute to choice. I partitioned the choice-aligned click train stimulus into bins on neutral-evidence trials (50Hz/50Hz), and determined on average, how well the balance of left and right clicks in each bin predicted the same trial's choice. Data are shown for our human and rat populations in figure 3.6a-b.

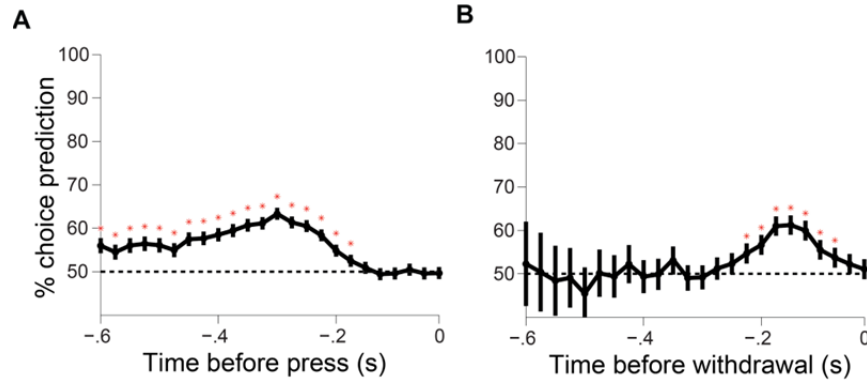


Figure 3.6: Evidence in the final moments of sampling does not predict choice. On each evidence-neutral (50Hz/50Hz) trial, click train pairs were aligned to choice and divided into 25ms bins. In each bin, the predicted choice was determined by the greater click count and compared with the subject's choice on the same trial. Bins where click count predicts choice are indicated with a red asterisk (where 95% binomial confidence intervals do not intersect with 50%). **A:** 5 Human subjects ($n=6,886$ evidence-neutral trials). Subjects listened to a pair of Poisson click streams and indicated their choice at time 0 as shown by pressing a response button. **B:** For the related rodent task discussed in Chapter 4, 6 rats ($n=3,262$ evidence-neutral trials). Rats listened to a pair of click streams and withdrew their snout from a port at time 0 as shown, stopping the stimulus. Some additional time elapsed while the rats moved to the response port, allowing them to make use of evidence in the sensory pipeline to reverse their initial judgments, while in contrast, the human button press instantly stopped the sound, and was a final commitment to choice. Our estimate of the earliest detectable time at which humans were totally committed to a response as measured by the last non-significant bin is 150ms. Due to using inter- port movement to register choices and our inability to distinguish between reversals and slow movement during this interval, the same estimate is not applicable to our rat study.

Our ROC analysis of the time course for which evidence predicts choice in section 5.4 further supports our estimate that evidence from the final 150ms does not contribute to choice. We chose to end the evidence window for computing β' at 150ms pre-choice. The β' measure is used in the following three chapters to describe the balance of experimentally delivered Poisson click evidence in analyses of choice. Since confidence reports were given following choice, the β measure (including the committed response period) was used for relating confidence to evidence. For analyses where sidedness is irrelevant, we show unsigned discriminability as $|\beta|$.

We wondered whether any part of this post-decision period of sampling represented voluntary commitment to a motor response preceding a physically imposed motor latency. If subjects were physically *unable* to inform choice using evidence in the final 150ms of the stimulus, trials

where they responded sooner than 150ms would have an equal probability of being correct or incorrect. Only three trials in our human data set had reaction times below 150ms (Figure 3.7a). Our subjects were verbally instructed to earn as much reward as possible, learning a balance between speed and accuracy – and prolonged sampling times are consistent with this strategy. They also knew they would have to provide confidence reports, and may have adjusted their sampling strategy to include a post-decision evidence collection period that *only* informed confidence after commitment to choice, in accordance with the post-decision confidence model reviewed in section 1.5.5. We wondered whether with a speed instruction and training, subjects could adjust their commitment time to make use of evidence closer to our detected button press response time. We tested three subjects on a new variant of our decision making task. After a random delay on each trial, subjects were presented with a binaural Poisson click stream with perfect evidence (100Hz/0Hz), and asked to indicate which side was clicking as quickly as possible. Trials were interleaved randomly between left and right. Perfect evidence was used to discourage strategic integration, and no confidence reports were requested. The time between the click train onset and the choice response would thus represent the summed physical latency of the button press response after choice, and the added delays of sensory processing and decision making, which studies in human and monkey vision have estimated at ~25-30ms (Bodelón, Fallah et al. 2007; Stanford, Shankar et al. 2010).

Since the three subjects had previously participated in the confidence task, they were first trained for 10 minutes in the speed condition with visual feedback; a green bar plot showed the mean of their *correct* reaction times, and the subjects were instructed to reduce the height of the bar as much as possible. During subsequent 10 minute probe sessions used for our analysis, no visual feedback was given. The reaction time distribution and conditional accuracy functions for the pooled probe sessions of both subjects are shown in figure 3.7b-c.

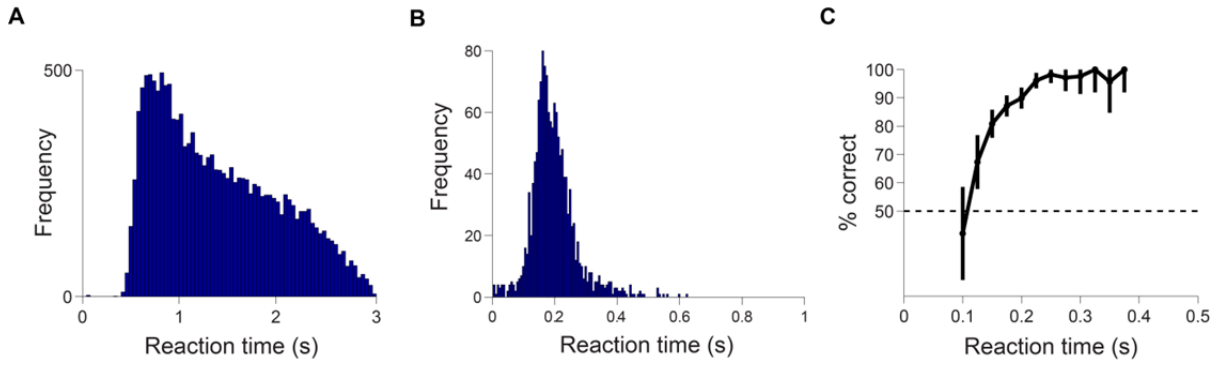


Figure 3.7: Subjects responding as soon as 125ms after stimulus onset discriminate above chance. **A:** Pooled reaction time distribution for 4 human subjects in the explicit-only confidence task (32,308 trials). Only three trials had reaction times below 150ms. **B:** Pooled reaction time distribution for three subjects performing the speed task with perfect evidence (1781 trials). **C:** Accuracy conditioned on reaction time is shown in 25ms bins. Bins with less than 20 data points were not shown. Confidence intervals show 95% binomial CI. This analysis shows that the minimum sensorimotor response latency for which evidence informed choice was 125ms. Accuracy rose above chance after 100ms, reaching the maximum distinguishable performance in this dataset by 250ms.

These results indicate that while subjects performing our confidence reporting task did not use information in the final 150ms before choice, this was not entirely due to a physical constraint on the speed of motor response and may have been modulated by the subject’s decision making strategy, alertness or other factors.

3.3.3. Click train evidence was the major determinant of subject choices

Subjects use click-train evidence to guide choice.

Since perceived evidence is the primary determinant of confidence as defined in equation 3.1, we sought to determine the degree to which our delivered evidence (and not other factors) influenced each choice. Psychometric functions, shown in Figure 3.8a-d, were plotted using our experienced evidence discriminability measure, β' (section 3.3.2), and fit with Weibull functions using maximum likelihood (see methods). Accuracy varied continuously as a function of graded discriminability in each of our subjects (Figure 3.8a-d). To compute post-hoc discrimination thresholds for each subject, I computed the strength of experienced evidence that resulted in 80% accuracy using the fitted Weibull function. Discrimination thresholds for subjects (indicated as “e80” in Figure 3.8 insets) were

comparable. The difference between our lowest and highest scoring subjects was 0.023 differential units of $|\beta'|$, equivalent to 2.3 extra differential clicks per second (of 100Hz summed between the two channels).

Our four subjects completed 15,834 trials in the explicit-only task variant. Mean choice accuracies for Subjects H1-H4 averaged across the experimental dataset were comparable: 81.1%, 81.7%, 83.8%, 84.5%. To ensure that subjects discriminated reliably, we computed the time course of discrimination accuracy within and across sessions (Figure 3.8e-f). Subjects discriminated consistently throughout the experiment, and any fatigue within sessions did not affect accuracy.

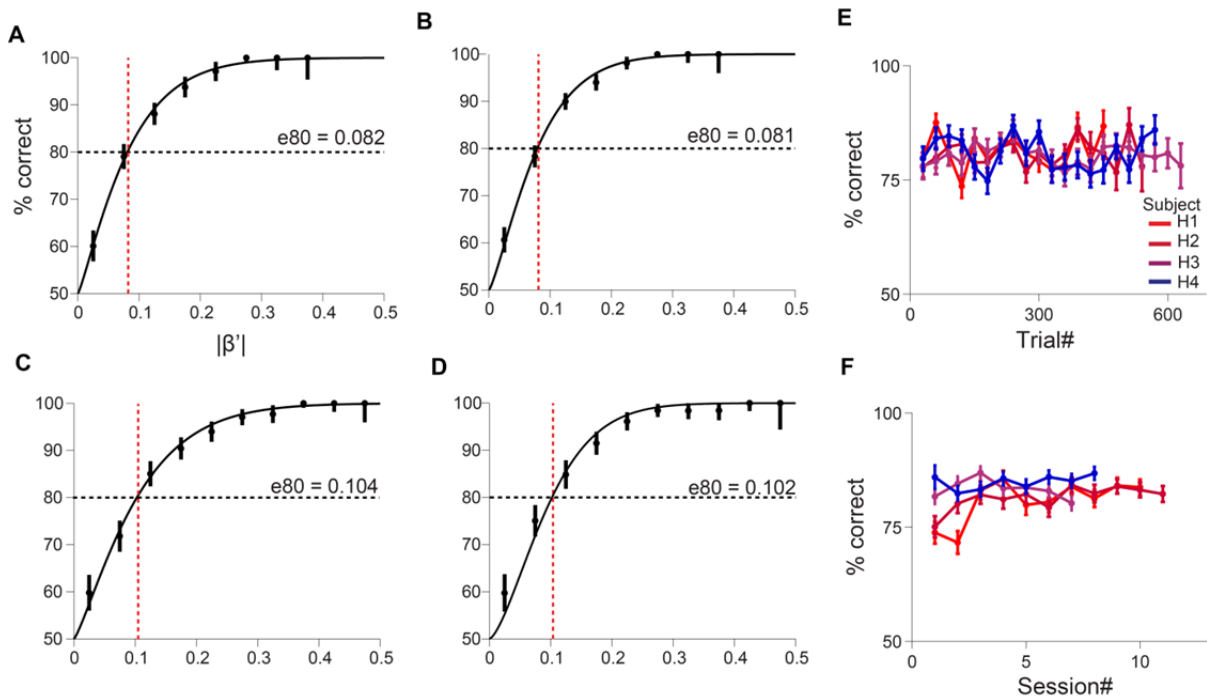


Figure 3.8: Discrimination accuracy is stable and varies as a function of our discriminability measure. A-D: Psychometric functions are shown for subjects H1-H4 respectively, in the explicit-only task variant. Points and error bars show means and 95% binomial confidence intervals for choice accuracy in each evidence bin. Solid lines are Weibull functions fitted to the bin means. The text inset shows the discriminability for which each subject performed at 80% accuracy (e_{80}), and the same value is indicated by a dotted red line. A dotted black line shows 80% choice accuracy. Points were not shown if fewer than 20 measurements contributed to the mean. **E:** Mean accuracy is shown averaged across sessions in 30 trial bins. Throughout sessions, accuracy for *all* subjects remained within 7 percentage points of 80%, and did not decrease towards the end of sessions as would be

expected if subjects fatigued. **F:** Accuracy is shown for each session following the two calibration sessions (section 2.1.2) for each subject.

Our delivered evidence dominates choice, despite small side and reward-history biases.

Especially since click stream pairs were directional and our subjects were not given clinical hearing tests prior to our study, we wondered whether they exhibited side bias. Probability of left choice was plotted as a function of signed β' (experienced click balance) to determine the extent of side bias in our subjects. I fit these data with a logistic regression model (see methods) and using the fitted function, evaluated excess left choice percentage when subjects were given perfectly neutral evidence (indicated in Figure 3.9a-d insets as “B”). To determine the significance of the bias for each subject, I computed 1000 bootstrap replications of the bias statistic B by re-sampling sessions of trials with replacement from our data, and re-fitted the logistic model to each bootstrap sample. 95% confidence interval boundaries were provided by the 2.5th and 97.5th percentiles of bootstrap distribution, and displayed as red dotted lines in Figure 3.4a-e. Two subjects (H1 and H2) did not have a statistically significant side bias. One subjects had a slight but significant bias of $+2.2 \pm 1.8\%$. The remaining subject had a larger, significant side bias of -8.9% .

To evaluate the relative contribution of reward history to each choice, I computed separate psychometric functions for trials following correct left and right feedback, and for trials following left and right errors. Trials following correct feedback are shown with logistic regression fits in Figure 3.9e-h, and error feedback in Figure 3.9i-l. Bootstrap distributions of the B statistic for each curve pair were computed as had been done previously to test side bias. From the resulting bootstrap distributions, the overlap of 95% confidence intervals was used to determine whether subjects responded differently after feedback on left and right response trials. Three of four subjects went left more often after being rewarded for going left, though the magnitude of this effect was less than 10%. Biases after errors were mostly non-significant, and effect direction was inconsistent among subjects.

We noted in a recent study employing a similar Poisson click discrimination task in humans and rats (Brunton, Botvinick et al. 2013), that rats often adopted a strategy where the first click on

either side disproportionally biased choice. While the authors played the first click of each stream in stereo to discourage this strategy, we did not take this precaution in our human study in order to maintain true Poisson evidence, which simplified reverse correlation analyses. To determine whether our subjects used a first-click discrimination strategy, I analyzed neutral-evidence trials (50Hz/50Hz), and predicted that choice on each trial would be the side to click first. For the four subjects included in the present analysis, this strategy predicted $46.9 \pm 3.4\%$, $49.5 \pm 2.8\%$, $55.4 \pm 3.1\%$ and $59.2 \pm 3.3\%$ of actual choices (errors are 95% binomial confidence intervals). However, we tested one additional subject in the task, for who first click predicted choice on $74.1 \pm 1.6\%$ of trials, implying that this subject was not using evidence in the way we intended. How robust the normative model patterns are with respect to this different regime of evidence evaluation is examined separately as a case study in appendix III.

In summary, we controlled only one factor as our intended evidence – the rate balance of a pair of Poisson click trains, and needed to determine to what degree subjects used this evidence to drive choice. We observed a strong dependence of choice on our parameterization of experienced click train evidence, and comparatively minimal effects of choice bias, click order or history dependence, indicating that our parameterization of evidence strength, β' , would provide significant insight into decision confidence.

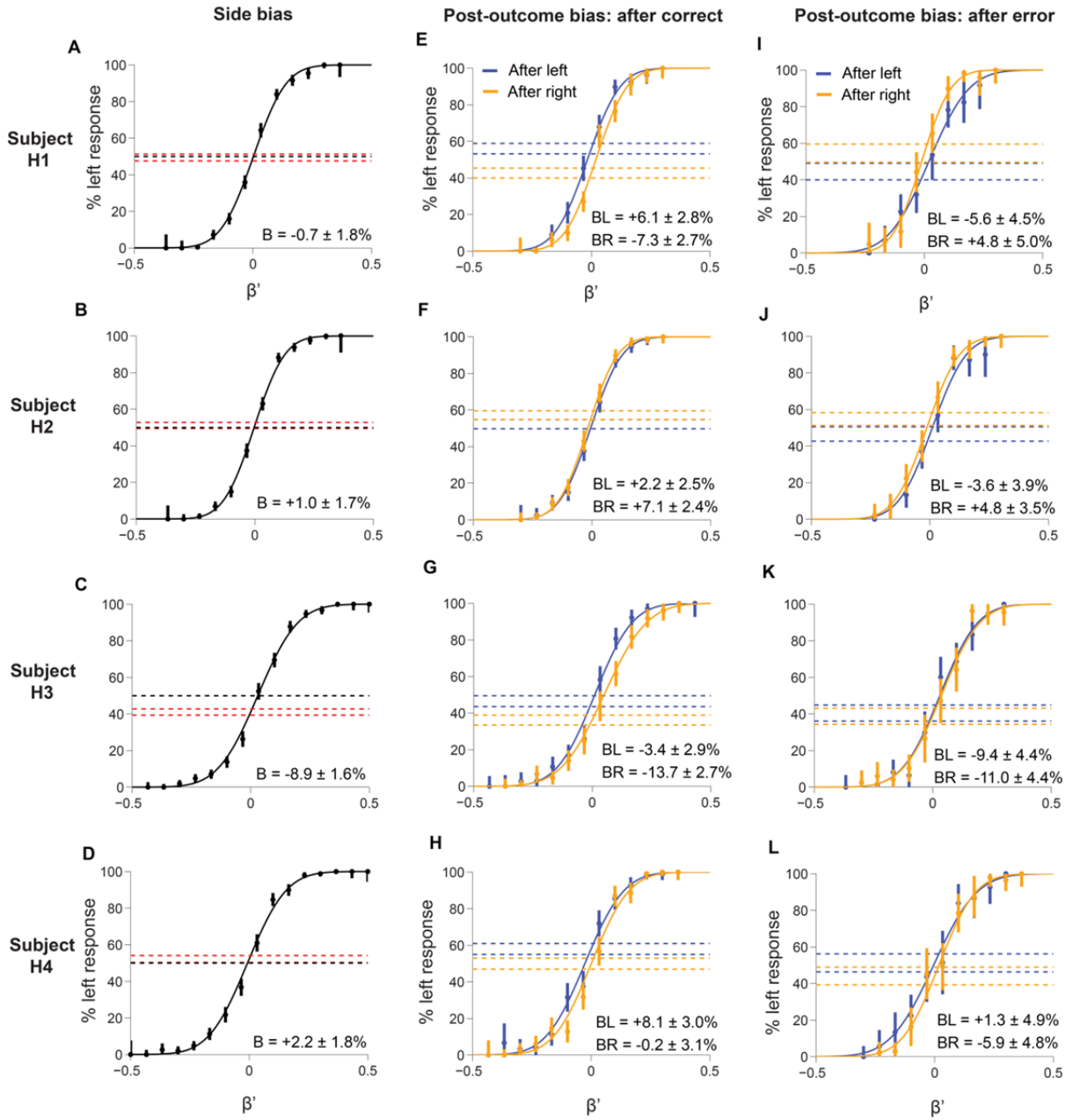


Figure 3.9: Click stream evidence primarily informed choice, despite small side and reward-history biases. Three analyses are shown on separate rows for each subject who participated in the explicit-only task. **A-D:** Sided psychometric functions are shown for all subjects. Points and error bars are means and 95% binomial confidence intervals for the percentage of left choices in each evidence bin. Solid lines are logistic regression fits to the same points. Net side bias (B), computed as the unsigned difference between 50%-left and the percent left for neutral evidence (determined by the curve fit), is printed on the text inset of each plot with 95% confidence intervals. Pairs of dashed red lines are drawn surrounding the same confidence interval, and a dashed black line indicates 50% left choice. Choice bias is minimal for most subjects and the direction of the effect is not consistent (as might be expected if the headphones were not calibrated to have equal click intensity). **E-H:** Choice

bias is shown for trials following left and right choices that were rewarded. Separate psychometric curves were fit for post-left and post-right conditions, and bias was computed as for panels A-D. Blue curves indicate trials after left choices, and orange curves indicate trials following right choices. Though it is a small effect in most cases, bias on the subsequent trial generally favors the side rewarded previously. **I-L:** Choice bias following error trials. Effect size is small and bias direction is not consistent. Axis labels for each plot are indicated on the top row.

3.3.4. Human subjects adopted stable evidence sampling strategies

Subjects discriminating based on temporally structured stimuli often adopt a sampling strategy that reflects the quality of evidence. If evidence is strong, subjects choose quickly, conserving time. If evidence is weak, subjects sample for longer, averaging out stimulus and sensory noise to gain a performance advantage. The precise duration of sampling on a trial of a given discriminability may depend on the cost of time for each subject, and their instantaneous estimate of the value of continued sampling (Drugowitsch, Moreno-Bote et al. 2012). Subjects were provided with a range of individual evidence strengths to compensate for different discriminability thresholds (H1-H4: 0.082 0.081 0.104 and 0.102, see Figure 3.8e). The mean reaction times of subjects provided with strong evidence were shorter than for subjects provided with weaker evidence (means in Figure 3.10a), and these individual differences largely persisted across sessions (Figure 3.10d). We show in Figure 3.10b that our subjects spent extra time sampling more difficult stimuli, consistent with a decision policy employing speed / accuracy trade-off (Palmer, Huk et al. 2005). To quantify the direction of this relationship, I fit linear trend lines to the data with least-squares regression, and computed bootstrap confidence intervals for the regression slopes by resampling trials with replacement and re-computing the linear fit for each bootstrap sample. All trends were significantly negative ($p < 0.05$). I also show that average performance decreases with continued sampling time in Figure 3.10c, consistent with the hypothesis that weak evidence trials are preferentially sampled for longer.

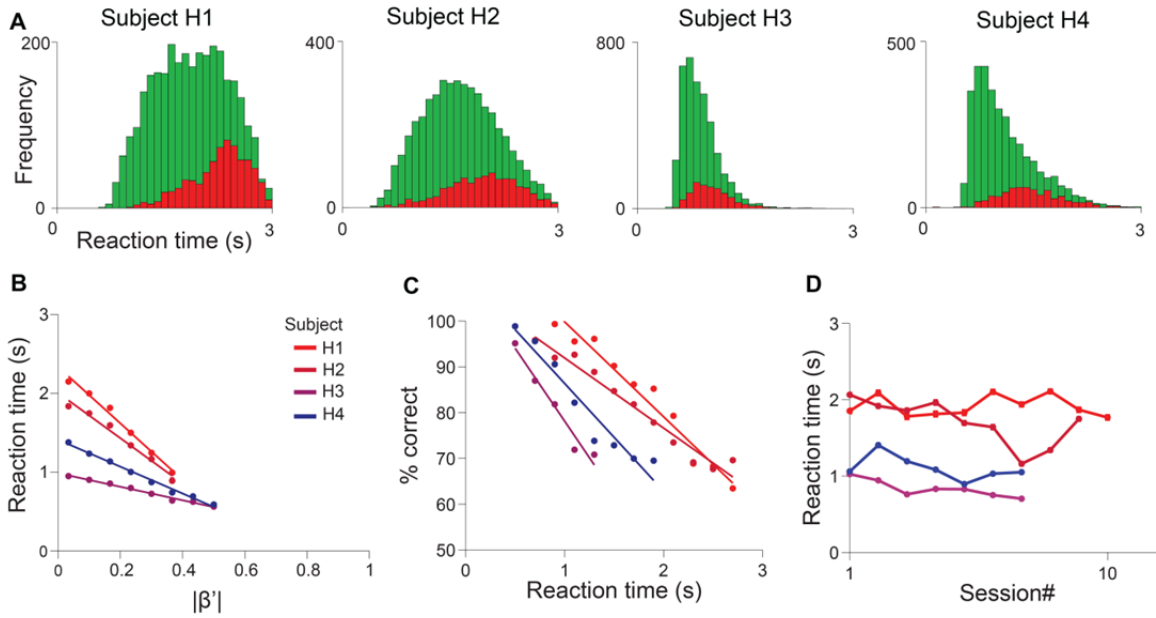


Figure 3.10: Subjects sampled evidence strategically to economize time. **A:** Distributions of reaction times for subjects H1-H4 are shown for correct trials in green and for error trials in red. All subjects sampled for significantly longer on trials when they erred (see table 3.1). Subjects provided with strong evidence to compensate for a high discrimination threshold sampled for less time on average. Distributions appear skewed forwards in time, except for two cases where reaction times regularly approached the 3 second sampling time limit (Subjects H1 and H2). **B:** All subjects sampled for more time before responding on trials with weaker evidence. Trend lines were fit to data with least squares regression. Trend line slopes and slope errors are listed in Table 3.1, and were all significantly negative. **C:** For all subjects, trials with long reaction times had lower accuracy than short-RT trials. **D:** With respect to larger differences between subjects, mean reaction time was stable across sessions (though note subject H2), indicating that the means of distributions shown in Panel A characterized persistent response strategies.

Subject ID	Correct RT mean (s)	Error RT mean (s)	RT / $ \beta' $	%correct / RT
H1	1.84±0.02	2.23±0.03	-3.58±0.15	-21.1±2.0
H2	1.60±0.02	1.91±0.03	-2.88±0.15	-15.0±1.2
H3	0.81±0.01	0.96±0.02	-0.86±0.05	-31.8±6.1
H4	1.05±0.02	1.40±0.04	-1.70±0.06	-24.1±4.2

Table 3.1: Significance of relationships between reaction time, evidence and accuracy. The table shows means and 95% confidence intervals of correct and error RT distributions in figure 3.5a. All correct trial reaction time means are significantly smaller than error means. In the fourth and fifth column, the table shows slopes and bootstrapped slope errors of trend lines fitted to data in Figure 3.5b-d. Two trends are significant for every subject as determined by slopes that are negative within error: reaction time varies with the inverse of evidence, and accuracy varies with the inverse of reaction time. Errors shown are 95% CI.

The past four sections were intended to infer how well our measures of discriminability and outcome describe the process used by subjects to make choices. We determined that subjects almost exclusively used the evidence we provided to choose, and chose accurately in a way that was predictable from our measure of evidence strength. Furthermore, we observed patterns in reaction time consistent with the process of actively listening to temporally structured evidence to make choices. These results show that our measures of experimentally delivered evidence and choice outcome, required to observe normative model predictions, are good descriptors of the subject's behavior.

3.3.5. Verbal confidence reports reproduce predicted normative model patterns

Earlier, we derived three characteristic patterns of confidence reports that emerge from the normative model, and showed that these patterns are true of a familiar confidence metric: p-values used in statistical hypothesis testing. If confidence reports generated by the brain are approximations of statistical decision confidence, the same characteristic patterns should manifest in human verbal confidence reports. To test this, we observed our data in projections that should produce the three normative confidence patterns.

The first prediction is that confidence predicts choice accuracy. For each subject in our study, self-reported confidence indeed exhibited this pattern (Figure 3.11a-d). We report the confidence calibration metric (equation 1.2) and statistically significant positive regression slopes for this relationship in Table 3.2. Subjects were all well calibrated to their respective levels of accuracy, consistent with previously reported calibration values in sensory discrimination studies (Figure 1.1).

The second prediction is that confidence on correct trials is correlated with experimentally delivered evidence, while confidence on incorrect trials is anti-correlated with the same. We found that human confidence reports *do* show divergence of confidence with respect to evidence on correct and error trials (Figure 3.11e-h), characterized by significant positive regression slopes for correct trials and negative slopes for error trials in all subjects (Table 3.2).

The third pattern predicts that for a given level of discriminability of delivered evidence, confidence predicts accuracy. We found that this was true for each subject (Figure 3.11i-l).

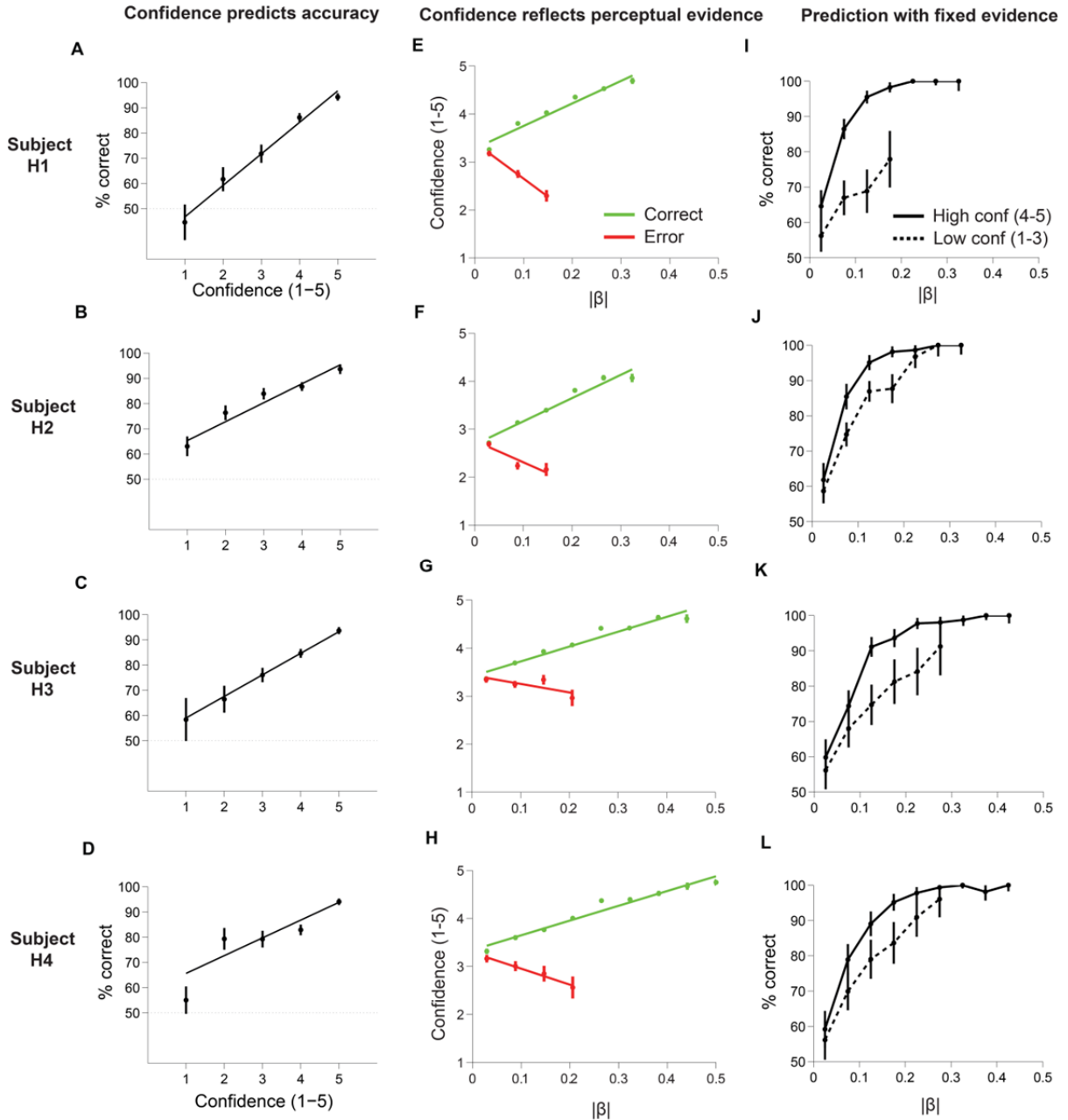


Figure 3.11: Confidence reports predict outcome and contain insight into perceptual evidence in all subjects. The patterns previously shown for statistical confidence in Figure 3.2 also characterized human verbal confidence reports. Data are shown in each row for four subjects who participated in the explicit-only confidence task variant. **A-D**: Calibration curves show that confidence is positively correlated with accuracy in each subject. Points are mean outcome for each reporting scale division, and solid lines were fit to the means with least squares regression. Slopes and

slope errors are reported in Table 3.2. Error bars indicate 95% binomial confidence intervals. **E-H:** Confidence reports are shown with respect to discriminability. The highest average confidence occurs in the presence of the strongest evidence when subjects are correct (shown in green), and the lowest confidence reports *also* occur in the presence of the strongest evidence - when subjects err (shown in red). Points and error bars are mean and standard error of confidence reports in each evidence bin. Solid lines were fit to the means with least squares regression, and slope errors are shown in Table 3.2 to indicate trend significance. **I-L:** When considering trials of the same discriminability, confidence still predicts accuracy. Solid curves are psychometric curves for high confidence trials, and dashed curves show low confidence. Errors are 95% Binomial CI. Axis labels are shown in the uppermost row.

Subject ID	Accuracy/Conf regression slope	Conf/Evidence correct slopes	Conf/Evidence error slopes	Calibration
H1	13.3±1.1	+4.6±0.3	-7.5±2.0	0.004
H2	7.1±1.5	+4.8±0.6	-4.4±3.4	0.020
H3	8.9±2.1	+3.0±0.3	-1.8±1.6	0.001
H4	7.2±1.0	+3.1±0.2	-3.1±2.1	0.013

Table 3.2: Significance of relationships between confidence, evidence and accuracy. The slopes and bootstrap slope errors are shown for trends depicted in Figure 3.6. The final column shows confidence calibration for each subject, computed with equation 1.2.

We have shown that human reports of confidence feelings have three robust properties that are characteristic of statistical confidence. We have shown that confidence predicts accuracy, consistent with a century of prior findings. For the first time in humans, we report that the lowest confidence is apparent in the presence of the greatest absolute provided evidence strength when a decision maker errs. In the next section, we will investigate some temporal properties of confidence reports that lend insight into how confidence reports are formed.

3.3.6. Verbal confidence properties are stable within and across sessions

Subjects exhibited confidence patterns despite a range of reaction time distributions and mean evidence strengths. We wondered whether subjects used the reporting scale consistently, and whether individual differences also existed in scale usage. Subjects were verbally instructed to use the full range of the reporting scale when providing reports on the first day of the experiment, and reminded again following the initial training sessions (section 2.1.1). Despite these instructions and similar

discrimination thresholds (Figure 3.8), subjects adopted a range of stable scale usage profiles. The frequency distribution of scale usage is shown for subjects H1-H4 in Figure 3.12a.

Since all of our analyses of confidence rely on data pooled across trials and sessions, we sought to determine the consistency of confidence reports. In Figure 3.12b-c, I show that mean confidence within and across sessions remains within one scale division for each subject. Figure 3.12d shows that subjects were well calibrated for individual sessions throughout the study, indicating that confidence was a good predictor of outcome (see section 1.1.1 for discussion of the confidence calibration statistic).

While the response scale was vertically oriented on the confidence response keypad, the decision component of our task had a sided response. We show in Figure 3.9 that some subjects had slight side biases. To determine whether confidence level varied systematically with side in individual subjects, we computed the ratio of mean left to right confidence across the study. For subjects H1 to H4, these values were: 0.96 ± 0.02 , 0.94 ± 0.02 , 0.94 ± 0.01 and 0.99 ± 0.02 (errors are bootstrapped 95% confidence intervals).

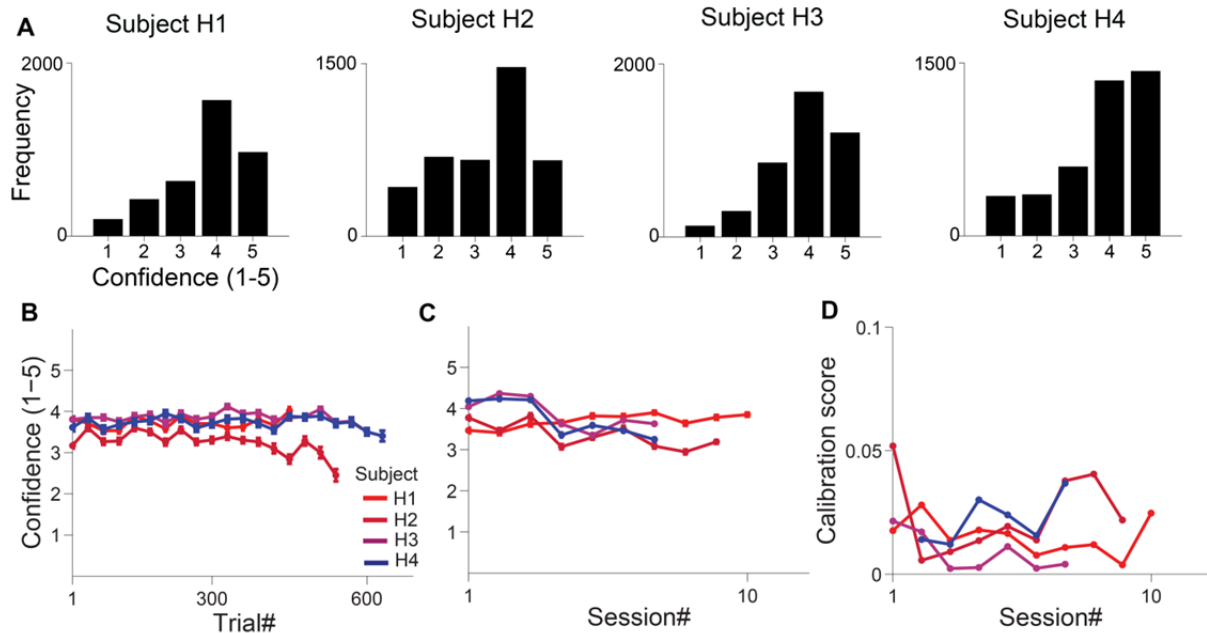


Figure 3.12: Subjects showed unique but consistent confidence levels. **A:** Frequency histograms of response scale usage show subject variability in response. **B:** Mean of confidence within a session. With the exception of subject H2, mean confidence scores did not decline towards the session end, consistent with flat session performance shown in Figure 3.3e. Data points are means and standard errors of 30-trial bins. **C:** Mean confidence for each session remained within one scale division for each subject throughout the experiment. **D:** Subject calibration scores showed that confidence persistently predicted choice outcome. Scores were typical of sensory discrimination tasks reported in confidence literature (Figure 1.1).

3.4. Human confidence in non-sensory decisions approximates statistical confidence

Our normative model defines confidence as the estimated probability of a correct outcome, given the evidence that was apparent to the decision maker. In the context of a perceptual discrimination task, our implicit assumption is that perceptual evidence is not equivalent to experimentally delivered evidence because experimentally delivered percept was distorted by noise. This noise could have originated in the sensory pipeline, or in the decision mechanism itself. In principle, the fundamental properties of confidence should not depend on *sensory* noise, since confidence judgments based on memory and general knowledge can be well calibrated (reviewed in section 1.1.1). We were especially interested to determine whether error trial confidence was anti-correlated with the strength of evidence in a regime where the sensory signal used to choose was likely not the origin of error.

We chose to evaluate normative model patterns in confidence judgments based on comparing quantities that are general knowledge instead of sensory percepts. I designed a geography quiz termed “Confography” which could be rapidly distributed by email for subjects to run in MATLAB. We collected data from 24 volunteer subjects in the Neuroscience research group at Cold Spring Harbor Laboratory. In this task, subjects were presented with the names of two countries and asked to indicate which country had a larger population within three seconds. They were then asked to report their confidence in their choice using the 5 point scale. To ensure that the subjects had perceived the country names correctly, 10% of trials were probe trials during which the subjects did not give confidence reports, and instead typed the names of the countries they had just been asked to compare

– a control for which subjects were never incorrect. A full account of the task contingencies is detailed in section 2.1.4.

Three subjects responded with less than 60% accuracy and were eliminated from the analysis. The remaining 21 subjects performed 2,550 trials – nine subjects performed 150-trial sessions while the remaining subjects performed 100-trial sessions. Subjects exceeded the three second response time limit on 2.2% of trials, and had an overall accuracy of 67.4%.

Since our measure of evidence strength depended on the assumption that true population statistics had been learned by the population prior to the study, I sought to eliminate countries whose populations were systematically misjudged with greater than chance probability, as determined by a 95% binomial confidence interval below 50% correct. This could occur if the population used heuristics to respond (for instance, “Australia is larger than Nepal so it must have a larger population”). In our dataset, no individual countries were misclassified at a probability greater than chance, however this does not preclude the possibility that a larger dataset would identify such effects.

I first introduce our results by establishing that our measure of city population difference predicted choice accuracy. I show that our data reproduces a fundamental effect reported previously in memory accession tasks: that subjects spend more time deliberating when they err. I then report our confidence findings.

3.4.1. Psychometric and chronometric performance on the Confography task

In designing the task, we assumed that the average subject had knowledge of the populations of world countries, and that countries with similar populations would be difficult to distinguish. Since country populations are never equal to zero, we used the absolute log ratio of each trial’s population pair as our measure of discriminability. Subjects determined the larger among country pairs with large population differences at above 90% accuracy, and performed at near chance when populations were similar (Figure 3.13a-b).

As with sensory discrimination tasks, subjects often respond more slowly when they err in memory access tasks (MacLeod and Nelson 1984). We found that this was slightly true of our data; pooled across subjects, correct trial reaction times averaged 1.87 ± 0.02 s, while error trials averaged 1.95 ± 0.03 . Distributions are shown in Figure 3.13c-d. Subjects spent slightly but significantly longer responding on difficult discriminations (Figure 3.13e-f), in contrast to the robust anti-correlation between reaction time and evidence we found in our sensory task (Figure 3.11b). A significant anti-correlation between reaction time and performance is shown in Figure 3.13e-f, consistent with drift diffusion model predictions for memory access (Ratcliff 1978). Since subjects responded on a wide array of hardware configurations, reaction time differences of less than a few hundred milliseconds were not considered reliable with this assay (indicated in section 2.1.4).

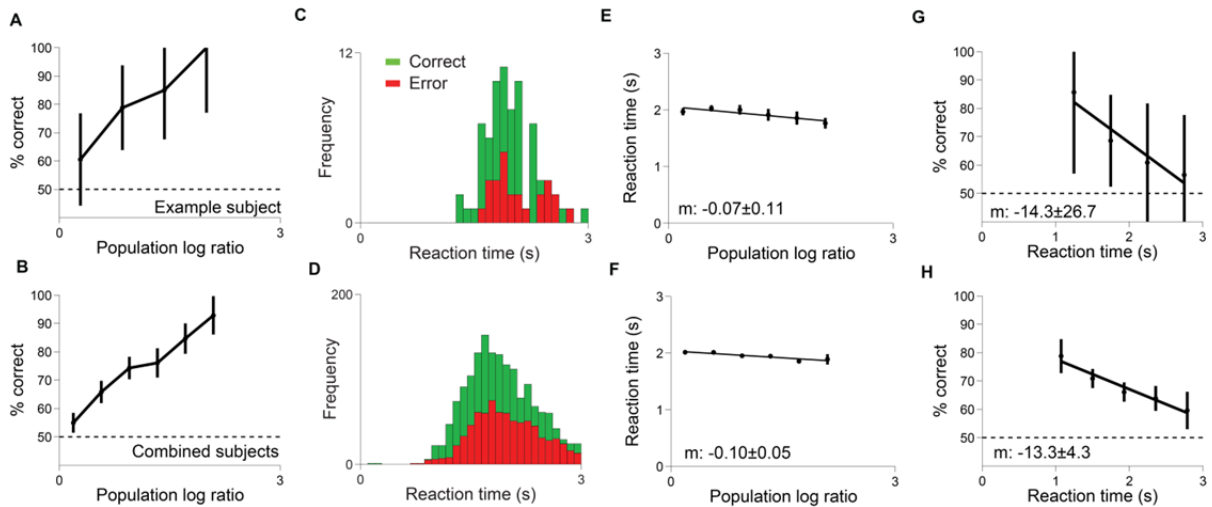


Figure 3.13: Population log ratio and RT measures support use of general knowledge to decide. The top row of panels shows data from a 100 trial session in a single example test subject. The bottom row shows combined data for our population of 21 test subjects. **A,B:** Subjects comparing two populations in memory were more accurate when the difference between them was larger. Performance means ranged from 54% in the most difficult bin of differences to 92% on the easiest. Errors are 95% binomial CI. **C,D:** Subjects took at least one second to respond, and took slightly longer to respond when they erred. **E,F:** Reaction times were slightly but significantly shorter on easy trials in our population, though our example subject did not show a significant trend. Slope errors are bootstrap 95% confidence intervals. **G,H:** Subjects were less accurate on trials with longer reaction times.

3.4.2. Verbal confidence in general knowledge reproduces normative model patterns

Subjects entered confidence reports following choice by pressing laterally oriented, numbered buttons on their computer keyboards as shown in Figure 2.9. Scale usage featured more use of low confidence scale divisions (Figure 3.14b) than in our perceptual study (Figure 3.12a), consistent with mean confidence calibrated to the difference between overall choice accuracies in the studies (67.4% vs 82.7%). However, individual subjects often had unique patterns of scale usage (Figure 3.14a).

With only ~100 trials per subject and a range of discriminability, analyses based on outcome in several categories were rarely significant in single subjects (Figure 3.14c). However, for our population, confidence strongly predicted outcome probability (Figure 3.14d), and the confidence calibration score of 0.04 was similar to previous reports in general knowledge and memory studies (Figure 1.1).

In both our example subject and our population, the lowest confidence occurred during the easiest incorrect decisions – despite the fact that the quantities being directly compared were memories of generally known quantities and not sensory percepts (Figure 3.14e-f). Correct choice confidence in the example subject and population was positively correlated with discriminability. The value and error of fitted linear regression slopes confirming the significance of the normative model's predicted trend directions are listed in Table 3.3. At fixed difficulties, confidence predicted accuracy, confirming our previous observations in sensory discrimination and statistical tests (Figure 3.14g-h).

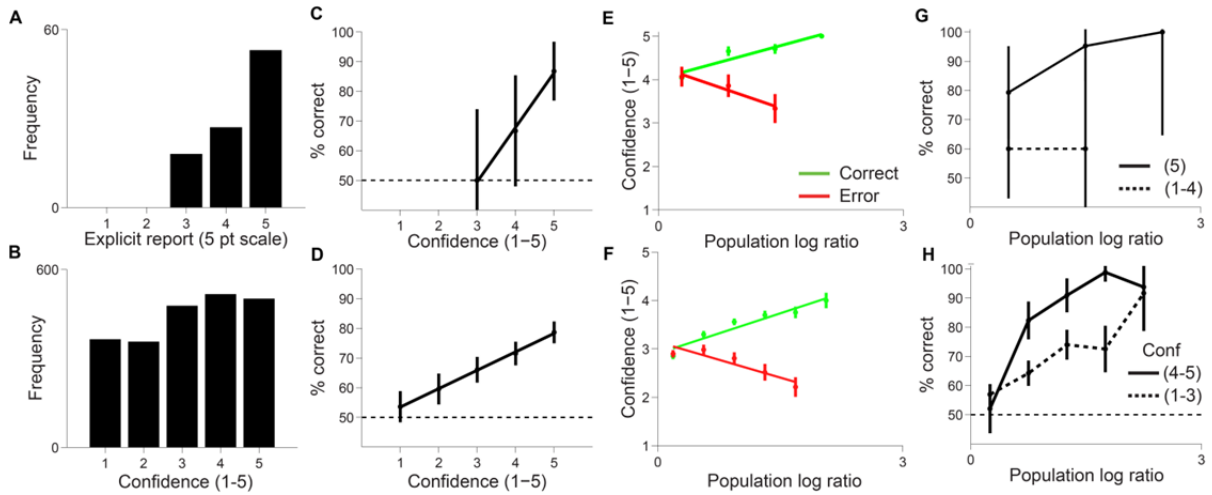


Figure 3.14: Normative model predictions describe confidence in general knowledge. The top row of panels shows data from the same example subject as shown in Figure 3.11, and the bottom row shows pooled data from 21 subjects (2550 trials). **A,B:** The population distributed their responses across the confidence reporting scale, while individual subjects had unique response patterns. **C,D:** Confidence reports were positively correlated with accuracy. Errors show 95% binomial CI. **E,F:** Correct and error confidence reports diverged with respect to discriminability. The lowest confidence occurred on strong evidence trials that were incorrect. Errors show 95% CI computed from SEM. **G,H:** At fixed discriminability, high and low confidence were distinguishable. Errors are 95% binomial confidence intervals.

Subject ID	Accuracy/Conf regression slope	Conf/Evidence correct slopes	Conf/Evidence error slopes	Calibration
GH15	+18.4±10.7	+0.49±0.03	-0.61±0.15	---
Population	+6.2±1.1	+0.41±0.02	-0.55±0.03	0.044

Table 3.3: Predicted normative model trend directions are significant for confidence in non-sensory choices. This table shows the slopes and bootstrap 95% confidence intervals for linear regression fits in Figure 3.9.

Since we observed a slight relationship between reaction time and evidence, we anticipated that this effect would be visible in confidence reports. In sensory and general knowledge tasks where subjects are free to sample evidence and enter choices at will, confidence is often (though not always) anti-correlated with reaction time (reviewed in section 1.1.3). In our sensory task, the magnitude of this effect was slight but significant (Figure 3.10a-d). In the Confography task, the anti-correlation for

both correct and incorrect trials was significant, though small (Figure 3.15a) and unlike in our sensory task, correct and error reaction times were indistinguishable in these bins.

The time between viewing the confidence prompt (section 2.1.4) and entering a confidence report was not distinguishable between correct and error (Figure 3.15b), but varied inversely with confidence in keeping with confidence reports from previous studies (also reviewed in section 1.1.3).

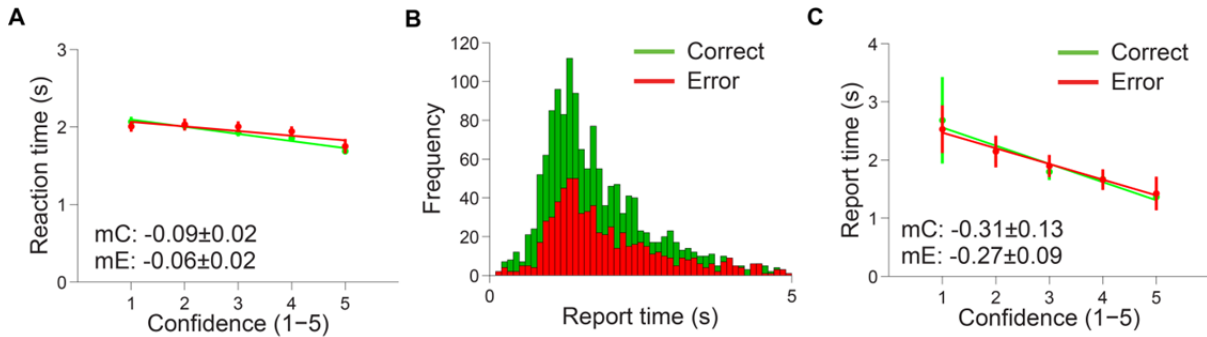


Figure 3.15: Temporal properties of confidence in general knowledge. **A:** Confidence is slightly but significantly anti-correlated with reaction time, though correct and error trials in each category do not have differentiable reaction times. **B:** The distribution of time taken to enter a confidence report following choice is skewed reflecting the absence of a time limit on reporting, and report latency is indistinguishable between correct and error trials. The correct trial report time mean was 1.87 ± 0.02 s, and for errors 1.95 ± 0.03 s where errors are 95% CI. **C:** The time to enter a confidence report varies with confidence in agreement with prior research (reviewed in section 1.1.3), consistent with a post-decisional deliberation process contributing to the reporting behavior.

We have shown that beyond statistical and sensory confidence, confidence in general knowledge exhibits normative model patterns. While the possibility remains that a corrupted mental representation of the sensory stimulus (country name text) informed memory access for choice, we show that the brain was able to recapitulate this information perfectly when probed, casting doubt on the likelihood that the confidence pattern we observed in evidence is *dependent* on sensory pipeline noise.

3.5. Conclusions

There is a strong precedent in literature detailing what decision confidence reports ought to look like. By necessity, our introduction of this precedent in Chapter 1 included studies from a century of literature in fields as diverse as Psychology, Economics, Neuroscience, Computer Science, Ethology, Geoscience and Medicine. With so many diverse reports informing the interpretation of our data, it is alarming that this dissertation is the first report of a new and fundamental property of confidence in humans: that confidence is *lowest* in the presence of the *strongest* absolute evidence strength when humans err.

We were able to derive this overlooked pattern from a perspective that clarifies confidence with a Bayesian definition, and makes strong predictions about the properties of pure confidence signals. We have shown that these predictions generalize from confidence in statistical hypothesis tests to people reporting confident feelings in two categorically different decision making tasks. These findings support our derived predictions as a general framework for identifying and understanding confidence measures.

Beyond statistical confidence and verbal confidence, these predicted patterns have been observed in the firing patterns of Orbitofrontal cortex neurons (Kepecs, Uchida et al. 2008), in a strategic time investment task performed by rodents. In the next chapter, we show that time investment behavior is correlated with verbal confidence in humans, and recapitulates normative patterns in human and rodent subjects, providing a way to understand *implicit* confidence within our framework as well as explicit confidence.

Chapter 4

Post-decision time investment provides access to decision confidence in two species

In forecasting accuracy, confidence provides a powerful advantage to decision makers managing risk, and has been regarded as an advanced metacognitive ability which may be unique to primates (Flavell 1979; Metcalfe and Shimamura 1994; Terrace 2004). A variety of tasks have been used to measure confidence in non-human primates, usually relying on binary decisions to either choose among options or to decline difficult choices for a small reward (reviewed in Chapter 1). More recently, an assay in rodents has shown that rats can strategically reinitiate trials after waiting for uncertain reward, providing a graded measure of investment (Kepecs, Uchida et al. 2008). How the human notion of confidence maps on to these behaviors is poorly understood. To bridge this conceptual gap, we developed a task that measures time investment in humans and rodents, and allows within-trial comparison of verbal and implicit confidence reports in humans. Using this task, we show that verbal confidence is strongly correlated with time investment. We also show that when considered with respect to accuracy and discriminability, time investment in both species is characterized by the normative model derived patterns expected of a confidence report. This crucial link between rodent behavior and the subjective feeling of confidence provides an opportunity to study the neural circuits underlying confidence with powerful molecular tools available only in rodent. More generally, our task provides a way to measure how confidence is used in situ to guide behavior with high resolution across species.

4.1. Normative confidence patterns characterize the time investment measure in human and rat

A convenient experimental assay to measure strategic investment of time in rodents was first described by Kepecs et al 2008. A subject registers a choice among two options with a motor action, placing the subject in a “waiting” state for that response. Following a random delay, the subject is rewarded if correct. On incorrect trials and on a small fraction of correct trials, no reward is delivered.

In theory, the subject invests time in the belief that they were correct, and this belief is proportional to the quality of evidence that informed the choice.

While investment of time is a tempting approximation of confidence, it is informed by several other contributing factors: the subject's cost function for time, and the utility of each expected reward. I begin this section by suggesting how each of these factors contributes to our measured investment of time, and how a pure confidence metric could be recovered in theory. I report two new adaptations of the time investment task for humans and rats (Figure 2.3), using auditory cues instead of odors. Using these task variants, I show that normative model predictions strongly govern the behavior of time investment in humans, and that this result generalizes across sensory modalities in rats.

4.1.1. Theoretical relationship of time investment to abstract confidence

The following equation describing the relationship between optimal time investment and confidence was generously contributed by Dr. Alex Koulakov, a primary investigator at Cold Spring Harbor, in support of our research:

$$TI = \tau * \log \frac{C * (1 - \tau * k)}{\tau * k * (1 - C)} \quad (4.1)$$

Where TI is time investment, C is confidence, τ is the reward probability decay constant defined by our exponential reward delay distribution, and k is the opportunity cost, defined as the average subjective reward value the animal experiences per unit of time while performing the task. The decision maker, having formed a mental representation of the distribution of reward delays, determines its optimal time investment as the time at which the cost of continued waiting exceeds the value of an expected reward, given its estimated likelihood.

Decision making organisms are likely to have different momentary opportunity costs – this variable is derived from both the average reward rate based on task contingencies, and the utility of an individual reward for the subject.

While equation 4.1 suggests a simple way of understanding the relationship between confidence and time investment in an ideal observer, several of these parameters may manifest differently in organisms. An animal's subjective estimate of time is imperfect (Eagleman, Peter et al. 2005), adding uncertainty to the observable effect of τ . Furthermore, animals behaving for liquid reward become sated, decreasing the opportunity cost throughout sessions in unpredictable ways. To directly infer confidence from time investment, in principle, requires a complex task design that compares time investment across otherwise identical conditions with different reward amounts and different reward delay distributions. To provide an initial characterization of whether time investment exhibits the patterns of a statistical confidence measure across species, we used nearly identical, fixed trial contingencies for humans and rats.

4.1.2. Human subjects perform the intended time investment task

We tested five human subjects in the time investment task (section 2.1.2). In exploring the relationship of time investment to accuracy and evidence, we first sought to ensure that our subjects were using the evidence provided to guide their choices. In Figure 4.1a, we show that choice varied as a function of graded click balance. Unlike our earlier study, two subjects had side biases above 10% for neutral-evidence stimuli (H3 and H6). Subjects showed a wide range of mean reaction times (Figure 4.1b), and always spent significantly more time sampling evidence on error trials (Table 4.1). The variable sampling durations were correlated with discriminability thresholds (Figure 4.1 insets); subjects who were poor discriminators were given stronger evidence to match performance among subjects, and needed to sample for less time. Performance for each subject across sessions is shown in Figure 4.1c, indicating that individuals were at a performance plateau and did not significantly improve between sessions (with the exception of subject H6) – an effect we would expect to contribute to variability in confidence if present. Reaction time means were also stable across the experiment (Figure 4.1d), indicating that sampling strategies were consistent. All subjects sampled for longer on difficult trials (Figure 4.1e) and were less accurate on trials with long sampling times,

consistent with strategic sampling of evidence to economize time. These patterns in performance and reaction time show the same trends in the same directions as for human subjects performing the explicit-only task variant (Figures 3.8, 3.9, 3.10), indicating that the fundamental process of evaluating evidence and choosing was not impaired by subsequent time investment.

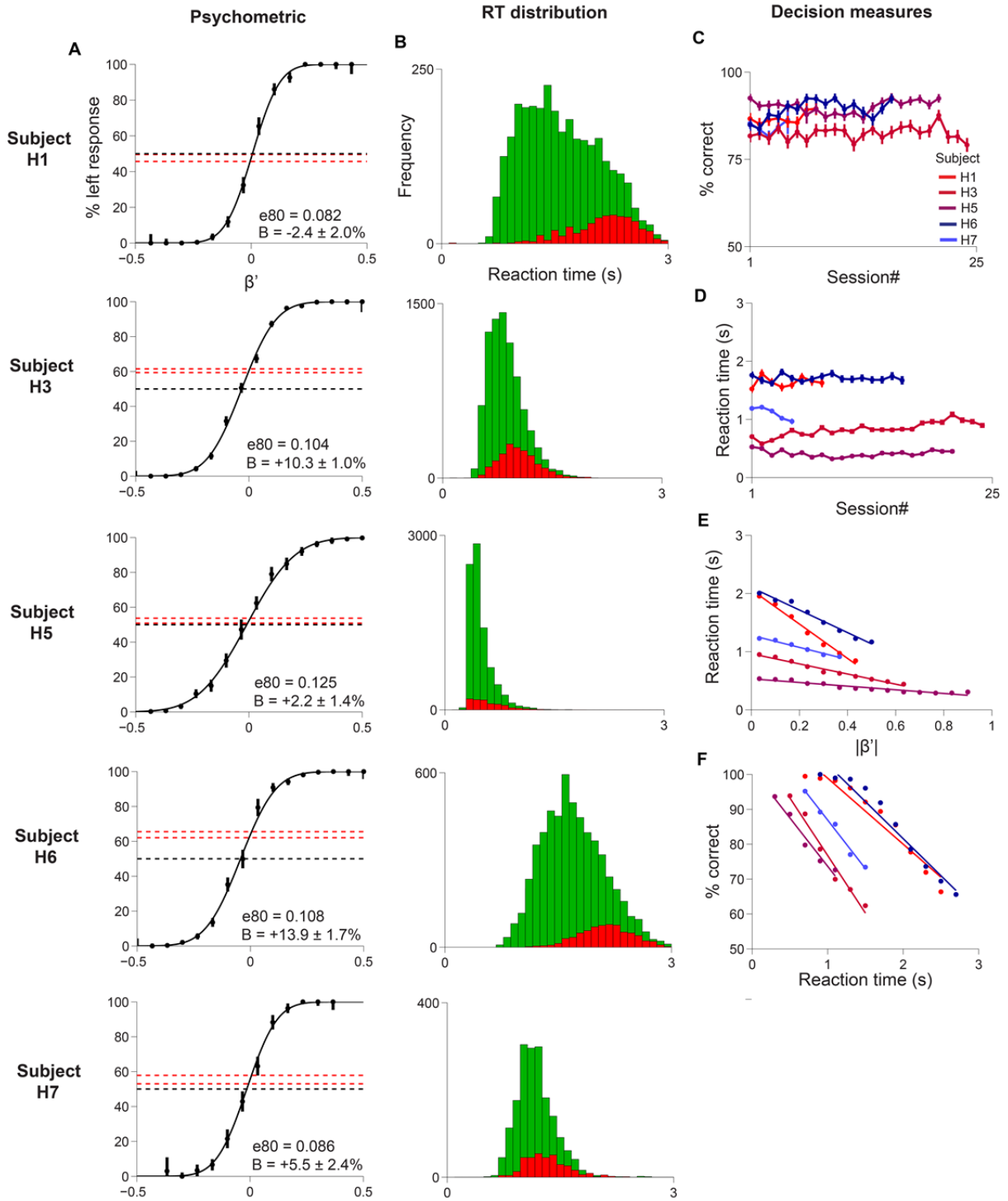


Figure 4.1: Subjects sampled Poisson click evidence to maximize performance. **A:** Sided psychometric functions are shown for all subjects. Points and error bars are means and 95% binomial confidence intervals for the percentage of left choices in each evidence bin. Solid lines are logistic regression fits to the same points. Net side bias (B), computed as the unsigned difference between 50%-left and the percent left for neutral evidence (determined by the curve fit), is printed on the text

inset of each plot with 95% confidence intervals. Pairs of dashed red lines are drawn surrounding the same confidence interval, and a dashed black line indicates 50% left choice. Side bias is minimal for most subjects (but see H3 and H6). **B:** Distributions of reaction times for all subjects are shown for correct trials in green and for error trials in red. All subjects sampled for significantly longer on error trials (see table 4.1). **C:** Mean choice accuracy is shown for each session. Errors are 95% binomial CI. **D:** Mean reaction time is shown for each session. Errors show SEM. **E:** Reaction time is shown with respect to discriminability for each subject. Subjects sampled for longer on difficult trials. **F:** Reaction time is shown with respect to choice accuracy for each subject. Subjects were less accurate on trials where they sampled for longer.

Subject ID	Correct RT mean (s)	Error RT mean (s)	RT / $ \beta' $	%correct / RT
H1	1.54±0.02	2.10±0.04	-2.95±0.09	-18.9±2.3
H3	0.80±0.01	0.97±0.01	-0.87±0.02	-32.9±4.0
H5	0.40±0.01	0.51±0.01	-0.31±0.01	-27.7±7.8
H6	1.65±0.01	2.09±0.02	-1.94±0.08	-21.2±2.2
H7	1.10±0.01	1.26±0.03	-1.03±0.09	-27.8±6.2

Table 4.1: Significance of relationships between reaction time, evidence and accuracy. The table shows means and 95% confidence intervals of correct and error RT distributions in figure 4.1b. All correct trial reaction time means are significantly smaller than error means. In the fourth and fifth column, the table shows slopes and bootstrapped slope errors of RT trend lines fitted to data in Figure 4.1e-f. Two trends are significant for every subject as determined by slopes that are negative within error: reaction time varies with the inverse of evidence, and accuracy varies with the inverse of reaction time. Errors shown are 95% CI.

4.1.3. Human time investment produces patterns expected of a confidence measure

All subjects learned to hold the response button and wait for reward, but individuals showed a range of willingness to invest time in unlikely returns (i.e. Subject H1 vs. Subject H7 in Figure 4.2 and table 4.2). Reward omission trials occurred at 10% frequency, allowing us to compare investment on correct and error trials. Subjects invested more time when they were correct (Figure 4.2), indicating that the time investment measure contained predictive information about outcome.

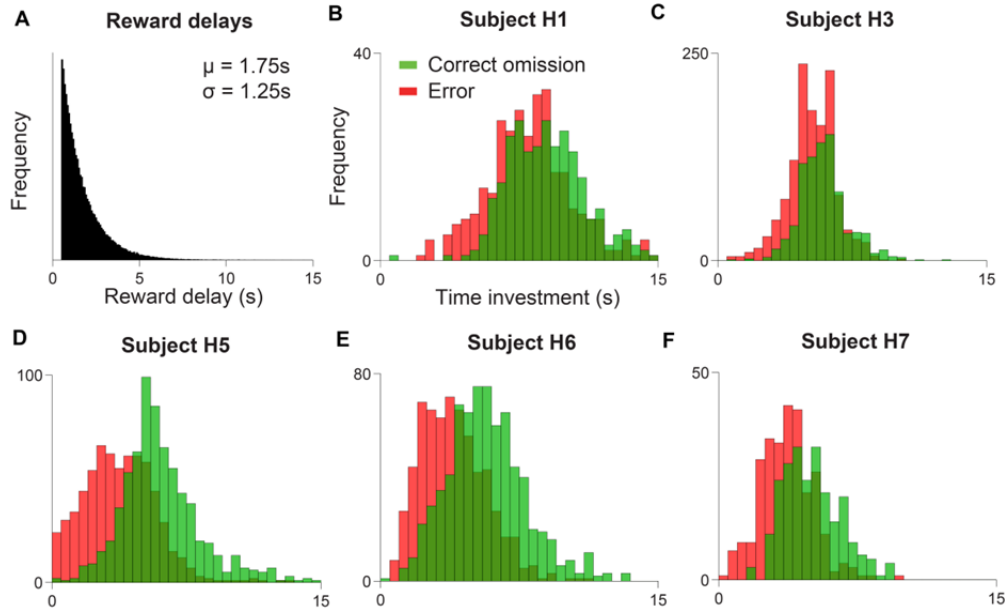


Figure 4.2: Human subjects invested liberally and invested more time when they were correct. **A:** the same exponential distribution of reward delays was used for each subject. The average delay was 1.75s, and only 2.7% of delays were greater than 5 seconds. As for our rats, delays below 0.5s were omitted to discourage impulsive abort responses. **B-F:** Time investment distribution for each subject. Correct reward omission trials are shown in green, and error trials in red. Time investment means are shown in Table 4.2; subjects invested significantly more time on correct trials. Investments were liberal in the sense that while the probability of reward for waiting longer than 10 seconds was virtually zero (panel A), human investments frequently exceeded this value.

Subject ID	Correct TI mean (s)	Error TI mean (s)
H1	9.13±0.36	8.02±0.32
H3	5.91±0.09	5.27±0.07
H5	6.24±0.19	3.83±0.17
H6	5.90±0.09	5.27±0.07
H7	5.03±0.20	3.70±0.17

Table 4.2: Human subjects invested more time in responses that were correct. Means and 95% confidence intervals (computed as $SEM \times 1.96$) are listed for the reaction time distributions shown in Figure 4.2. All correct time investment means were significantly greater than error means.

Time investment predicted accuracy, though unlike explicit reports and normative model predictions, the relationship was non-linear as shown for the average of all subjects in Figure 4.3a and for individual subjects in Figure 4.4a. Instead of a fitting a linear slope to measure this correlation, we show in Figure 4.3a that the range of time investments spans the range of accuracies as a monotonic function from chance to nearly perfect.

Time investment, like verbal confidence, was lowest on high discriminability error trials and highest on high discriminability correct trials (Figure 4.3b for the population and 4.4b for individual subjects). For fixed discriminability, time investment still predicted outcome (Figure 4.3c and 4.4c). Thus, the *directions* of the trends predicted from normative confidence are strongly apparent in human time investment decisions.

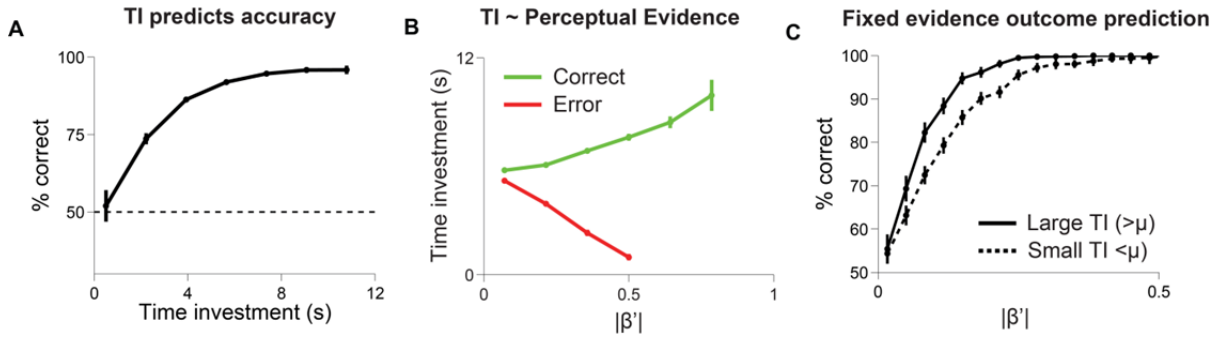


Figure 4.3: Human time investments resemble normative model predictions for confidence. Time investment is shown averaged across 5 subjects (34,446 trials) in the three projections where our normative model predictions would be apparent for a confidence measure. **A:** Time investment is nonlinearly related to accuracy, but spans the entire range of accuracies from chance to nearly perfect. Error bars show 95% binomial CI. **B:** Time investment is highest for high discriminability correct trials, and lowest for high discriminability errors. Time investment becomes less indistinguishable between correct and error trials as discriminability is reduced. Error bars show 95% CI. **C:** At fixed discriminability, subjects were more accurate on trials where they invested more time. Errors show 95% binomial CI.

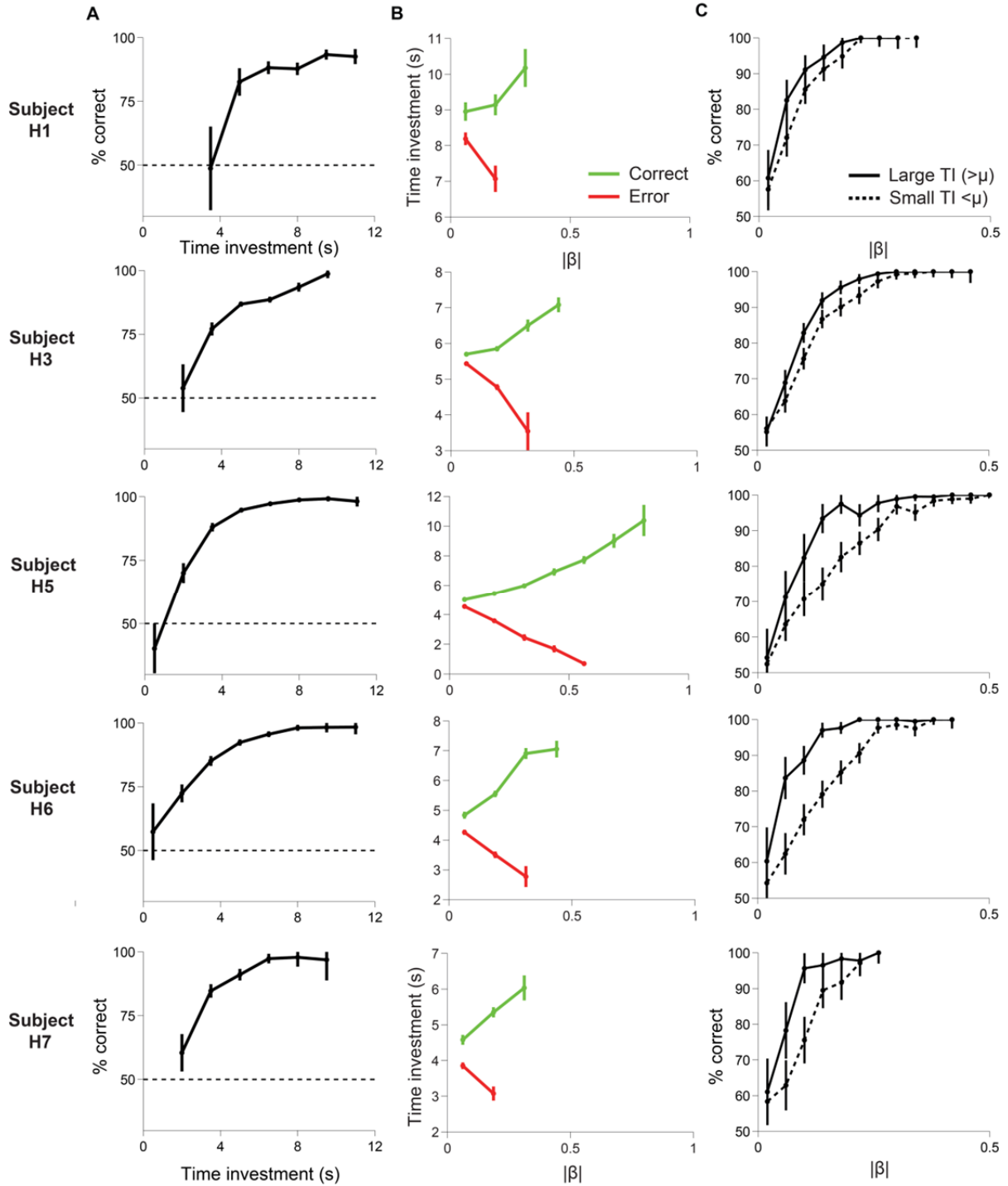


Figure 4.4: Human time investments for individual subjects resemble normative patterns. The patterns shown in Figure 4.3 for the population are shown for individual subjects. **A:** Time investment is nonlinearly related to accuracy. **B:** Time investment is highest for high discriminability correct trials, and lowest for high discriminability error trials. **C:** At fixed discriminability, subjects were more accurate on trials where they invested more time. All error bars are the same as for Figure 4.3.

4.2. Human time investment is strongly correlated with verbal confidence

In the previous section, we showed that time investment contains four trends in directions predicted by the normative model. To further explore the link between time investment and confidence, we acquired verbal confidence reports following time investment on a fraction of trials (section 2.1.2). We reasoned that if time investment was informed by confidence (section 4.1.1), it should be closely correlated with explicit confidence reports. I show that on trials where both explicit and implicit measures were captured, they are closely correlated (Figure 4.5a) regardless of outcome (Figure 4.5b).

In our understanding of the relationship between confidence and time investment (equation 4.1), time investment is derived from confidence – and for the subsequent verbal report, *confidence* is then consciously interrogated and reported by our subjects. While the correlation of these two within-trial measures is compelling, we wondered whether it could be explained by a reversal of the dependent relationship between the two measures – that when prompted for a verbal report, subjects instead infer their confidence from their own investment behavior. In Figure 4.5c, we show that verbal confidence is not simply a function of time investment; for a fixed time investment, confidence varies as a function of evidence. We also show that post-investment explicit reports are linear with respect to accuracy (Figure 4.5d), similar to pre-choice reports and unlike time investment, further supporting the idea that verbal reports were independently determined. The correlation between time investment and verbal confidence was consistently high throughout the experiment (Figure 4.5f) and was high in each subject (Figure 4.5g).

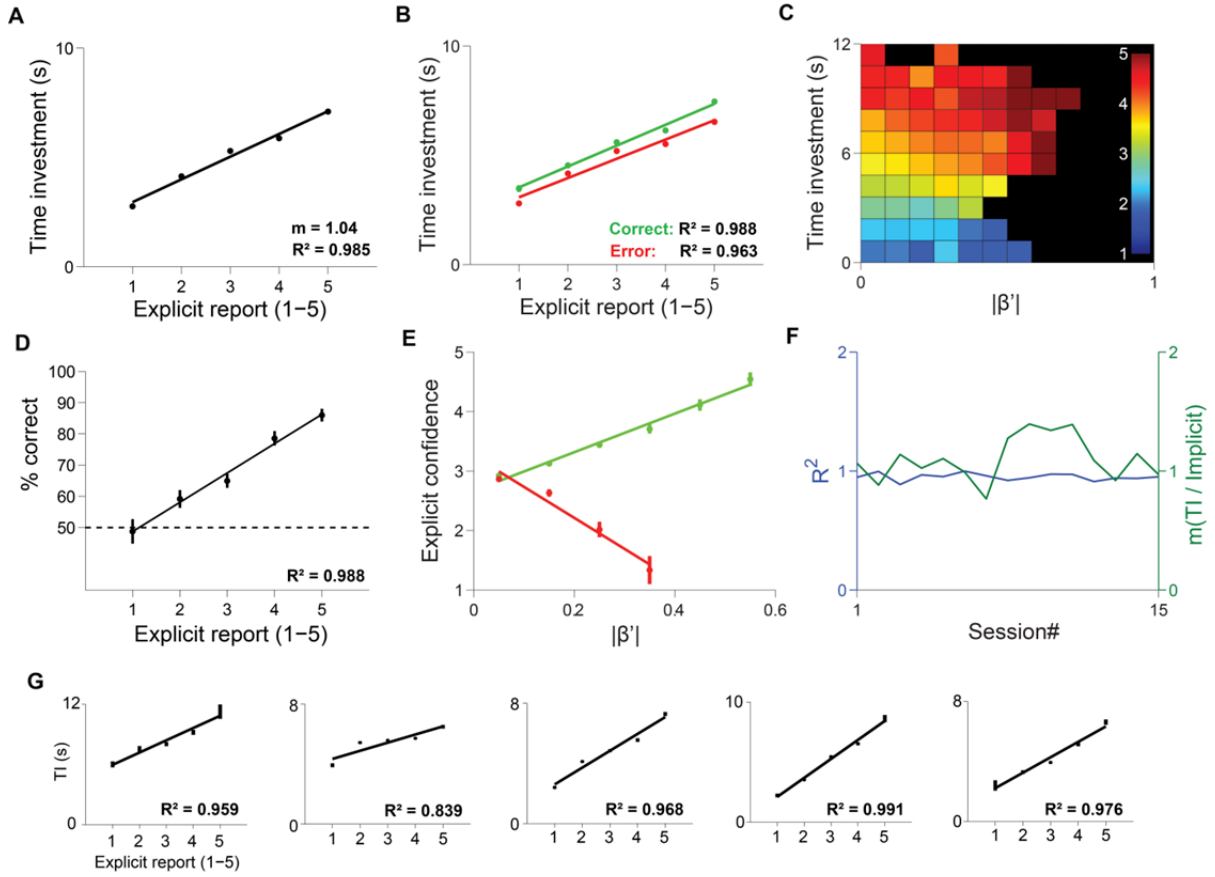


Figure 4.5: Time investments are highly correlated with subsequent verbal confidence reports. With the exception of panel G, panels show data averaged across 5 subjects (34,446 trials). **A:** For each division on the 5-point confidence reporting scale, mean time investment is shown as a single point. The slope and coefficient of determination for a line fitted with least-squares regression (solid black line) are shown in the figure inset. Verbal confidence is strongly and linearly correlated time investment. Error bars indicate 95% confidence intervals. **B:** Time investment and verbal confidence are correlated regardless of trial outcome. **C:** The interrelationship between time investment (ordinate), discriminability (abscissa) and verbal confidence (indicated by inset color scale-bar). For a given time investment, verbal confidence varies with discriminability. **D,E:** Contrary to time investments, post-investment confidence reports show nearly linear trends with respect to choice accuracy and evidence. **F:** For the correlation of time investment with verbal confidence (panel A), the slope was positive and the linear fit was a good description of the data throughout the experiment. **G:** For each subject, time investment and verbal reports were closely correlated. From left to right, subjects H1,H3,H5,H6 and H7 are shown.

We have shown that time investment by human subjects produced the trends expected of a statistical confidence report, though without preserving linearity between confidence and accuracy, consistent with the hypothesis that time investment is a measure derived from confidence. We showed

that these explicit and implicit measures were strongly correlated. To infer that the correlation was not caused by subjects mapping the previous time investment to the explicit scale when asked for verbal confidence, we showed that confidence varies as a function of evidence for a given time investment. These results strongly suggest that time investment provides a way to infer the brain's confidence, in assays where the cost of time and opportunity are known.

4.3. Rat time investment produces patterns expected of a confidence measure

Since the time investment measure does not require verbal ability, it is in principle applicable to non-human animals. A previous study (Kepecs, Uchida et al. 2008) showed confidence-like patterns in time investment by rats discriminating odor mixtures. We wondered whether the use of confidence to guide time investment in rats would generalize across sensory modalities, for direct analogy with our results linking verbal reports to time investment.

To assess this, we trained five rats in the rodent variant of our implicit time investment task (section 2.1.3). In brief, rats were placed in a behavior box with three response ports on one wall, and trained to enter the center port to initiate a trial. After a random delay, the rat sampled two separate streams of Poisson clicks delivered from speakers behind the left and right walls. Once the rats determined the faster clicking side, they withdrew from the center port, instantly stopping the click streams, and entered the lateral reward port corresponding to their choice. In the reward port, rats waited for a randomly delayed drop of water. On 10% of trials, reward was omitted despite correct responses to measure correct trial time investment, and every incorrect trial yielded a time investment measure. No additional timeout was added to the punishment of fruitless investment on error trials.

While the same overall click rate (100Hz distributed between sides) was used for both species, several differences between this task design and the human variant were notable. To encourage long waits and discourage impulsive withdrawal, rats were given longer reward delays (mean = 2.30s vs. 1.75s in humans). Rats were also punished with a 3s time-out for withdrawing in the first 200ms of sampling, to discourage impulsive withdrawal, and were never given delays longer

than 8s. When humans pressed the response button, the stimulus was immediately stopped, indicating both the end of evidence sampling and the temporally aligned motor execution of choice. For rats, the stimulus was stopped on withdrawal from the center port, allowing them the opportunity to switch their original decisions based on information in the sensory pipeline before entering a response port. Finally, rats were rewarded with liquid drops following ~22 hours of water restriction – a reward whose subjective value was likely higher than 5 US cents offered to humans for correct responses.

Despite the same click rate sum (100Hz), rats discriminated much more poorly than humans given the same evidence balance (shallow psychometric function in Figure 4.6a). The difference in signal quality between acoustically isolated left and right headphone channels and speakers positioned outside a behavior box with solid, acoustically reflective walls may account for this discrepancy. Rats also sampled for very short durations on average before withdrawing from the stimulus port, averaging only 341ms of evidence collection. Despite such short reaction times, rats sampled for longer on difficult trials (Figure 4.6c), mirroring the extended processing of difficult evidence observed in our human subjects. Rats were also less accurate on trials where they sampled for longer durations (Figure 4.6d).

While reward delays were 500ms longer than humans on average (Figure 4.6e), rats waited for comparable average durations (mean = 5.01 ± 0.06 s, error = 95% CI). This observation is consistent with the possibility that the cost of time is lower for humans (who earn 5 cents per correct trial) than for rats (which are relieved of liquid deprivation when correct), though we did not design the experiment to explicitly estimate a cost function. Like humans, rats invested more time waiting for reward on correct trials than on error trials (Figure 4.6f).

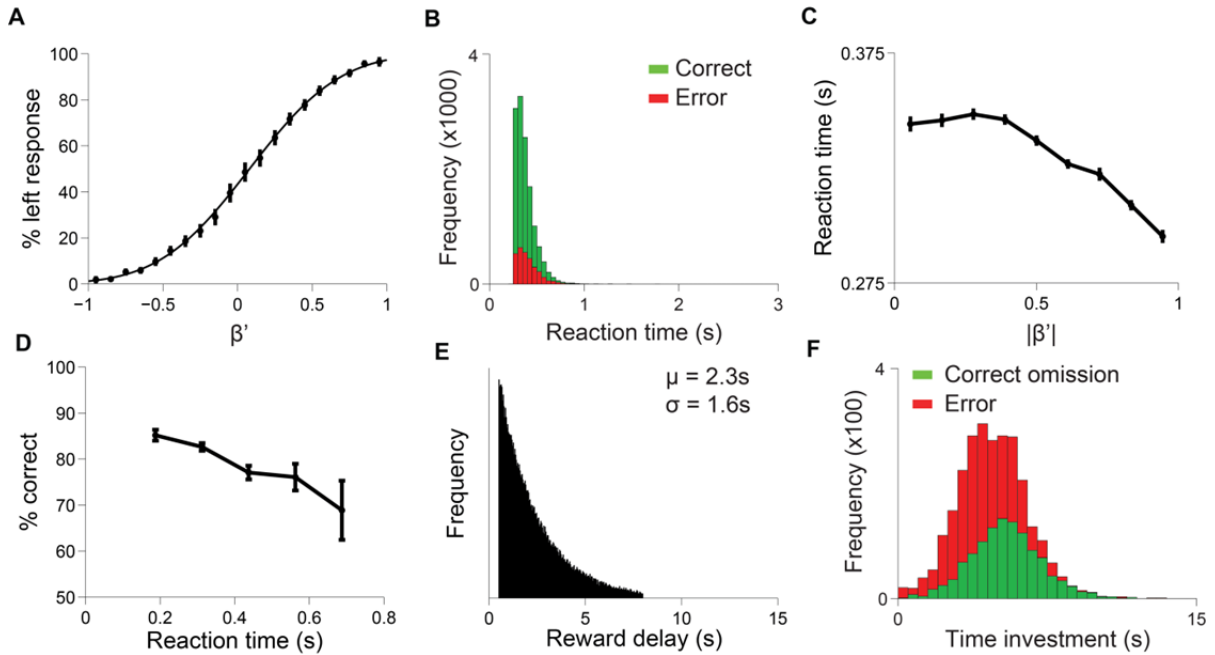


Figure 4.6: Rats perform the time investment task and invest more conservatively than humans. All data are averages across 5 rats who performed 21,499 trials. **A:** Psychometric function showing that chosen side varied continuously as a function of L/R click balance. **B:** Rats sampled for very short durations. Mean sampling time was significantly shorter for correct trials ($336 \pm 2ms$), than for error trials ($365 \pm 4ms$). **C:** Rats sampled for longer on trials with weaker evidence. Error bars indicate 95% CI. **D:** Rats were less accurate on trials with longer reaction times. Error bars are 95% binomial CI. **E:** Reward delays were longer for rats than for humans. While only 2.8% of delays exceeded 5 seconds for humans, 8.5% exceeded 5s in rats. **F:** Rats invested more time for correct trials than for error trials ($5.28 \pm 0.09s$ vs $4.88 \pm 0.06s$ respectively where errors are 95% CI). Despite a larger fraction of reward delays exceeding 5 seconds, rats spent less time waiting for reward than most humans (Table 4.2).

Time investment in rats performing olfactory discrimination had previously generated patterns consistent with our normative model predictions for confidence reports (Kepecs et al 2008). We found that rodent time investment contains these patterns despite a different sensory modality informing choice (Figure 4.7 A, C). As with our humans, the time investment conditional accuracy function was nonlinear. While the first two normative model predictions had been observed in rodents previously, the third prediction is reported in our data for the first time (Figure 4.7E).

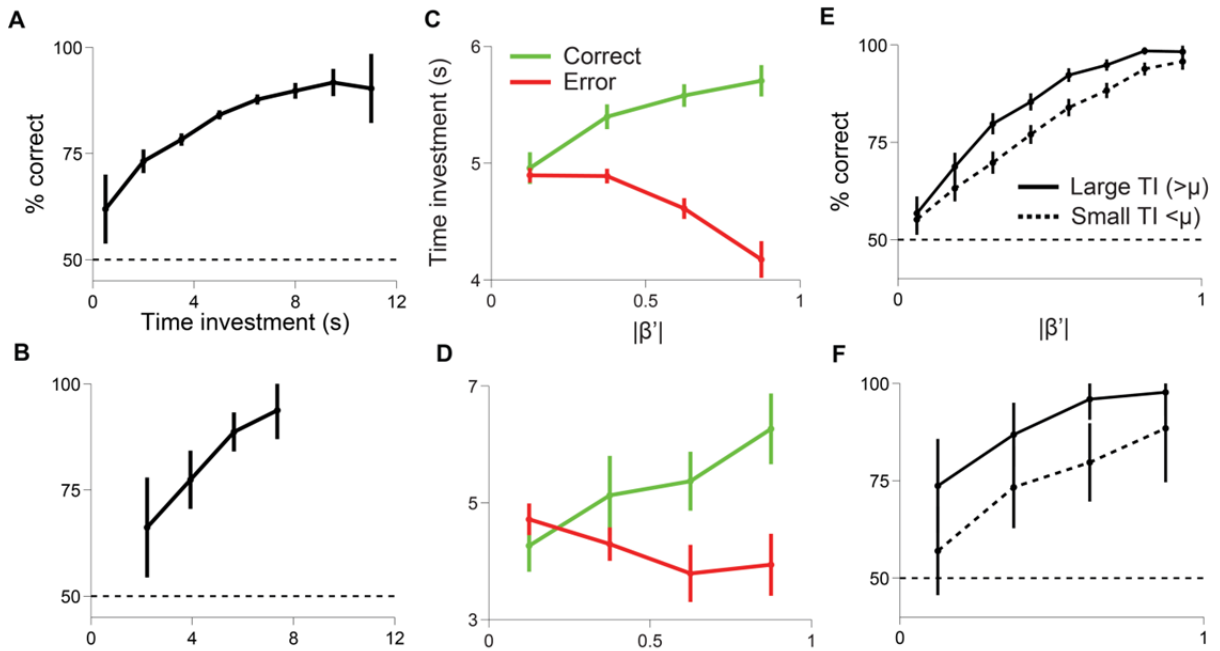


Figure 4.7: Rat time investments resemble normative model predictions for confidence. Panels A,C,E show pooled data from 21,499 trials in 5 rats performing the time investment task. Panels B,D,F show data from a single session of 522 trials (of which 31 were correct reward omissions and 102 were errors). The first two sessions of exposure to the final reward delay distribution and the first 20 trials of each session were omitted from the analysis. Sessions for which rats averaged below 90% accuracy on the easiest category of stimuli were also omitted. **A,B:** Time investment is nonlinearly related to accuracy, but spans the entire range of accuracies from chance to nearly perfect. Error bars show 95% binomial CI. **C,D:** Time investment is highest for high discriminability correct trials, and lowest for high discriminability error trials. Time investment becomes less indistinguishable between correct and error trials as discriminability is reduced. Error bars show 95% CI. **E,F:** At fixed discriminability, rats were more accurate on trials where they invested more time. Errors show 95% binomial CI.

4.4 Conclusions

The animal models most accessible to neurophysiology manipulations and measurements do not have linguistic abilities needed to provide explicit confidence reports, forcing most research on the neural correlates of confidence to use implicit measures. We have shown in two mammalian species, that strategic time investment displays three robust patterns expected of a confidence report, consistent with the possibility that a confidence metric is computed by the brain, and is used to guide time investment behavior.

We further show in humans that strategic time investment is correlated with verbal confidence when both are provided in series to measure the same prediction. This constitutes the first report directly comparing an explicit and implicit measure of confidence. Together with our previous findings, we have established an empirical basis for claims that time investment measured in rodents is indeed a confidence-derived behavior. By implication, rodents are capable of being confident – a finding that undermines previously argued positions on confidence as a higher order cognitive ability unique to primates.

There was a subtle difference between implicit and explicit reports in our data. Implicit reports mapped nonlinearly to accuracy, consistent with the hypothesis that time investment is a measure derived by combining confidence, expected reward probability and reward utility.

While we have established the presence of statistical confidence patterns in human and rodent time investment, how time investment is computed from abstract confidence algorithmically in different conditions remains an avenue for future behavioral research. Mechanistically, neurons in the orbitofrontal cortex have been shown to contain patterns indicative of a confidence signal (Kepecs, Uchida et al. 2008). With a validated behavioral confidence measure in rodents, continuing investigations may determine how this signal arises and contributes to strategic investment of an organism's time.

Chapter 5

The Opportunistic Coupled Accumulator (OCA) model of confidence

Behaving organisms are constantly faced with choices, and must collect and classify samples of evidence from the environment to inform them. In the previous two chapters, we determined that decision confidence is fundamentally a *statistical* computation performed by human decision makers on the evidence they used to decide. However, the form of the computation we proposed in equation 3.1 is abstract, and not equipped to describe how confidence is computed by an organism's brain as it sequentially samples noisy evidence in time.

Several computational models of choice and confidence have been proposed within sequential analysis and signal detection theory frameworks. These models use different algorithms to compute choice and confidence from evidence, implying different predictions about the features of evidence that inform choices, and about the general properties of confidence. To evaluate the fitness of these models, we designed a perceptual decision making task where choices and confidence were informed with high resolution, temporally structured evidence, similar to Brunton, Botvinick et al 2013. Unlike Brunton et al, subjects in our task were free to respond as soon as they had made a choice. Precise alignment of evidence history to the instant of a measured motor response could in principle, better enable us to identify different temporal patterns in the evidence preceding choice and confidence, if subject commitment to choice was precisely followed by the button press.

On each trial, subjects decided which of two Poisson click streams was faster on average, and subsequently reported their confidence in their choice on a 5-point scale (section 2.1.2). Unlike previous studies of confidence, this task afforded us access to a sub-millisecond precise temporal record of experimentally delivered evidence used by subjects to decide and generate a confidence report on each trial, allowing us to compare the patterns in evidence history left behind by humans to those produced by different sequential sampling models. Using previous models of choice and confidence, we attempted to simultaneously fit several important inter-relationships between five

trial-by-trial measures: choice outcome, discriminability, confidence, reaction time and evidence history. We found that each previous model failed to explain at least some aspects of the data. To resolve this discrepancy, we developed a new model: the Opportunistic Coupled Accumulator (OCA) model of choice and confidence. We anticipate that an algorithmic understanding of how confidence is computed by the brain in this simple decision task will inform future investigations of the more complex miscalculations of confidence that contribute to mental illness.

In the following five sections, we report properties of our choice and confidence data which cannot be explained by each of the five models of choice and confidence introduced in section 1.5. In the final section of this chapter, we provide a model that generates correct predictions and fits where previous models failed.

5.1 Methods

Our explicit-only confidence reporting task design is described in Chapter 2 and human task performance is characterized in Chapter 3. In brief, human subjects listened to a binaural stream of Poisson clicks on each trial and freely indicated which stream was clicking faster on average by pressing a left or right choice button when they felt they had made this determination. Each click was delivered in acoustic isolation and well above the human detection threshold for sound (70dB), presumably containing a clear (~100%) hypothesis about the faster clicking side. Delivering click streams and capturing button press events were accomplished with a custom device, affording a record of the precise times of clicks and button presses, and 50-microsecond precise control of evidence cessation after choice (section 2.1.5). After choosing, subjects were immediately prompted to report their confidence that their decision was correct on a 5-point pushbutton scale, where 1 indicated a random guess and 5 indicated high confidence. To vary the discriminability of each trial, the balance of left and right click rates was varied from 50Hz/50Hz (hardest) to ~65Hz/45Hz (easiest, precise value adjusted for each subject to match 90% performance). Subjects had up to three seconds to sample evidence, and forfeited the trial's reward if they responded after this interval. Five precise measures were captured to describe different aspects of the choice on each trial: the precise time

course of evidence sampled (click streams), the discriminability of evidence sampled (left/right click balance ratio), reaction time, outcome (correct or error) and verbal confidence (five point scale).

Task performance metrics and an analysis of patterns in confidence reports are provided in Chapter 3. In brief, we computed a measure of discriminability based on the balance of experienced clicks prior to choice (β'), and determined that subjects primarily used Poisson click evidence to drive decisions. We also found that reaction times were consistent with evidence integration, and that subjects performed with mean performance, reaction time and confidence consistent across sessions. Intriguingly, we determined that reports describing the human feeling of confidence exhibited a set of three patterns that characterize statistical confidence. We next sought to determine an algorithm which could produce these intriguing patterns in the human brain.

5.2 Reaction time models do not account for patterns in confidence reports

Previous studies of confidence have reported that under *most* circumstances, confidence is anti-correlated with reaction time (reviewed in section 1.1.3). To generate confidence from the drift diffusion model of choice (DDM), this relationship has been suggested to be causative – a decision maker using a drift diffusion algorithm would determine confidence by observing its own reaction times (reviewed in section 1.5.3).

To determine which patterns in confidence are predicted by the reaction time model, we fitted a drift diffusion model with four free parameters (threshold, integrator noise level, motor latency mean, motor latency variance, equation 5.1 with $\rho=0$) to the psychometric function and reaction time distribution of an example subject: subject H4 (see Chapter 3). For each trial, we set confidence equal to reaction time, providing a raw confidence measure. To map the model's raw confidence to the 1-5 scale, we determined the frequency with which subject H4 used each scale division, and divided the inverse raw confidence distribution into scale divisions by corresponding percentiles. Thus, a trial with a reaction time in the bottom fifth of the reaction time distribution would be scored a "5", while a reaction time in the top fifth would be scored a "1".

The psychometric curve produced by the fitted model fell within 95% binomial confidence intervals of the subject curve at each point (Figure 5.1a), and the model reaction time mean fell within one standard error of the subject (1.092 ± 0.004 s and 1.089 ± 0.007 s respectively, Figure 5.1b).

In the remainder of Figure 5.1, I show that while DDM can account for choice and reaction time, confidence reports show different patterns from the DDM model's reaction times. The reaction time model of confidence predicted that confidence is equivalent to reaction time, irrespective of whether the decision was correct (Figure 5.1c), and irrespective of discriminability (Figure 5.1d). Contrary to these predictions, human confidence is indeed differentiable between correct and error trials for each of our subjects (Figure 5.1e) and for a fixed reaction time, varies as a function of discriminability (Figure 5.1f). In human subjects, we observed that verbal confidence on error trials is anti-correlated with discriminability (Figure 5.1g, Chapter 3). For the same fit to our example subject, confidence on both correct *and error trials* was correlated with discriminability (Figure 5.1h). To determine whether the model was capable of producing an error-trial anti-correlation with different parameters, we evaluated model datasets of 20,000 trials, varying the two model parameters not related to the motor delay (noise level and threshold). Datasets were included if they produced choice accuracy within 10% of the mean of pooled human subjects (72%-92%), a reaction time mean within one standard deviation of the human mean (0.73-1.99s), and responded within the three-second sampling limit on 90% of trials. The slope of error trial confidence vs discriminability was never negative (Figure 5.1j).

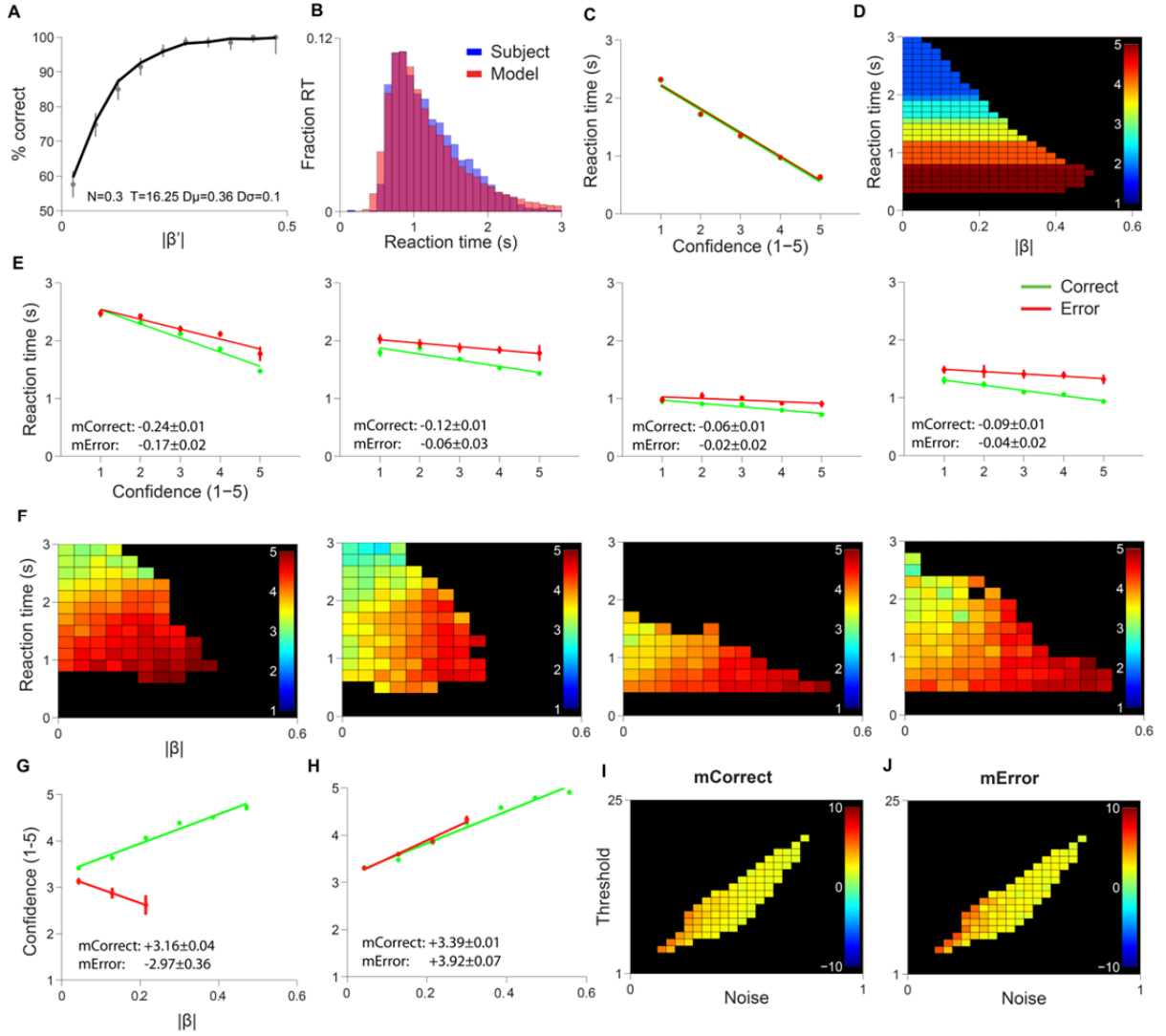


Figure 5.1: The reaction time model of confidence makes three incorrect predictions. A drift diffusion model was fitted to an example subject's data, to produce psychometric fits and reaction time means within error of the subject's data. **A:** Psychometric functions produced by an example test subject (gray points and 95% binomial) and the fitted drift diffusion model (solid black line). Model parameters are shown in text inset. **B:** Reaction time distributions produced by the test subject (blue) and the fitted model (red). **C:** For the fitted model, confidence for correct (green) and error (red) trials is strongly anti-correlated with reaction time, but is not differentiable between correct and error choices. **D:** Confidence (inset) is shown for the model with respect to both reaction time and discriminability. Discriminability is shown as the β measure (equation 2.1) – the ratio of the difference and sum of experienced left and right click counts. For a given reaction time, confidence does not vary as a function of evidence. **E:** Mean reaction times and 95% confidence intervals computed at each division in the confidence scale are shown for correct and error trials in subjects H1-H4 respectively. To quantify the significance of trends, slopes and bootstrapped 95% confidence intervals are shown in text insets. Confidence on both correct and error trials show significant anti-correlation with reaction time. Error trials have longer reaction times than correct trials of the same

confidence. **F**: Relationship between reaction time, discriminability and confidence is shown as a heat map. Scale inset shows confidence scale (Dark red = 5, Dark blue = 1). Within reaction time bins, confidence varies strongly as a function of evidence, a pattern that would not emerge if confidence were solely determined by reaction time. **G**: For the example subject (also shown in Figure 3.5), confidence is correlated with discriminability on correct trials, and anti-correlated on error trials in accordance with normative model predictions. Points are subject data. Errors are 95% CI. Solid lines are least squares linear regression fits. For each fit, text inset shows slope and bootstrap error. **H**: The model's confidence is correlated with discriminability on correct trials *and error trials*, violating normative model predictions. **I,J**: For correct and error trials respectively, model threshold and noise level were varied to determine whether the reaction model can produce error trial confidence that is anti-correlated with discriminability while producing reasonable reaction times and choice accuracies as described in text. This was not true for any set of these two parameters.

The reaction time model of confidence thus incorrectly predicts three robust patterns in human confidence which are evident in our data. While reaction time was significantly anti-correlated with verbal confidence in each subject (Figure 5.1E), we conclude that this is only a correlation, and not the algorithm used by the brain to compute confidence.

5.3 Variance models fail because variance in evidence is unrelated to confidence

Models of choice with relative stopping rules (like DDM) collect evidence for hypotheses A and B until evidence sum A exceeds evidence sum B by a fixed threshold. For this reason, they cannot use the difference in evidence collected for and against choice as a confidence measure, as can models using an absolute threshold (like race and other accumulator models). The reaction time model was one possible solution to this problem, especially plausible due to the oft reported anti-correlation between reaction time and confidence (Figure 5.1, but see section 1.1.3). An alternative algorithm recently proposed by Yeung and Summerfield (2012) is to derive confidence from a choice process with a relative stopping rule using a metacognitive process that explicitly computes the variance of the time course of perceptual evidence (reviewed in section 1.5.4). The model was intended to solve the task of identifying which “sensory hypothesis” was true, in a continuous space of possible sensory hypotheses, and thus departs from the two-choice drift diffusion framework while retaining relevance to sequential sampling problems. As a brief example of how their model computes

confidence, consider classification of a face in the continuous spectrum between male and female. Each noisy sensory sample of the face contains a hypothesis about maleness (e.g 30% male, 95% male). As more samples are collected, the mean of the sample distribution converges towards the true value, and the variance of the measurements provides confidence in the mean. Unlike accumulator and drift diffusion models, this model requires that the brain maintain a memory for the entire collection of samples in order to compute variance.

Earlier, we noted that acoustic isolation and high signal to noise (~70dB clicks) afforded us a regime where each click was assumed to contain a near-100% hypothesis supporting its side (but see Brunton, Botvinick et al 2013 for exceptions). This binary evidence afforded us the opportunity to determine whether variability in evidence, apart from discriminability, can account for confidence.

To quantify the variability of evidence, we computed two measures, the coefficient of variation (CV) for the entire click train leading to the choice, and also the local variation, known as CV2 (Holt, Softky et al. 1996). CV is the standard deviation to mean ratio of a trial's click intervals, while CV2 is a related measure that is less sensitive to changes in click rate as might be expected from a choice process that filters for momentarily elevated or suppressed rates and measures local variation in evidence. The mean CV2 was within error of 1, indicating that the average click train was Poisson ($CV2=0.999\pm0.001$, error = 95% CI). However, individual click trains were sub-Poisson or supra-Poisson, ranging in CV2 from 0.784 to 1.231, and indicating a range of click regularity on individual trials (Figure 5.2a,d). For chosen and anti-chosen click trains, confidence scale divisions did not have statistically distinguishable regularity of evidence by either measure (Figure 5.2b-c, e-f).

While confidence is unrelated to regularity of *delivered* evidence within each stream, a metacognitive process explicitly approximating the variance of a noisy internal decision variable may be less sensitive to the precise time course of evidence. To determine whether regularity of the decision variable's movement in a drift diffusion process could produce normative confidence patterns with Poisson evidence, we tested the standard deviation of momentary changes in decision variable value across its history as a confidence measure. For the most direct path, shown in bright

red (Figure 5.2g), the decision variable moves by exactly $-1 \cdot \text{Threshold}/RT$ units of evidence at each time step, and the standard deviation of the resulting vector of movements is 0, indicating the highest confidence. An indirect path (shown in dark red), would advance inconsistently and have higher variance in momentary movement, indicating lower confidence. For human subjects, confidence is strongly correlated with choice accuracy (Chapters 1, 3). However, this internal variance measure does not meaningfully predict accuracy for the fitted drift diffusion model used in section 5.2 (Figure 5.2h) or for the same model with a range of reasonable noise and threshold parameter values (Figure 5.2i). Slopes for the relationship of confidence scale divisions to choice accuracy in our subjects ranged from 7.1-13.3 (table 3.2).

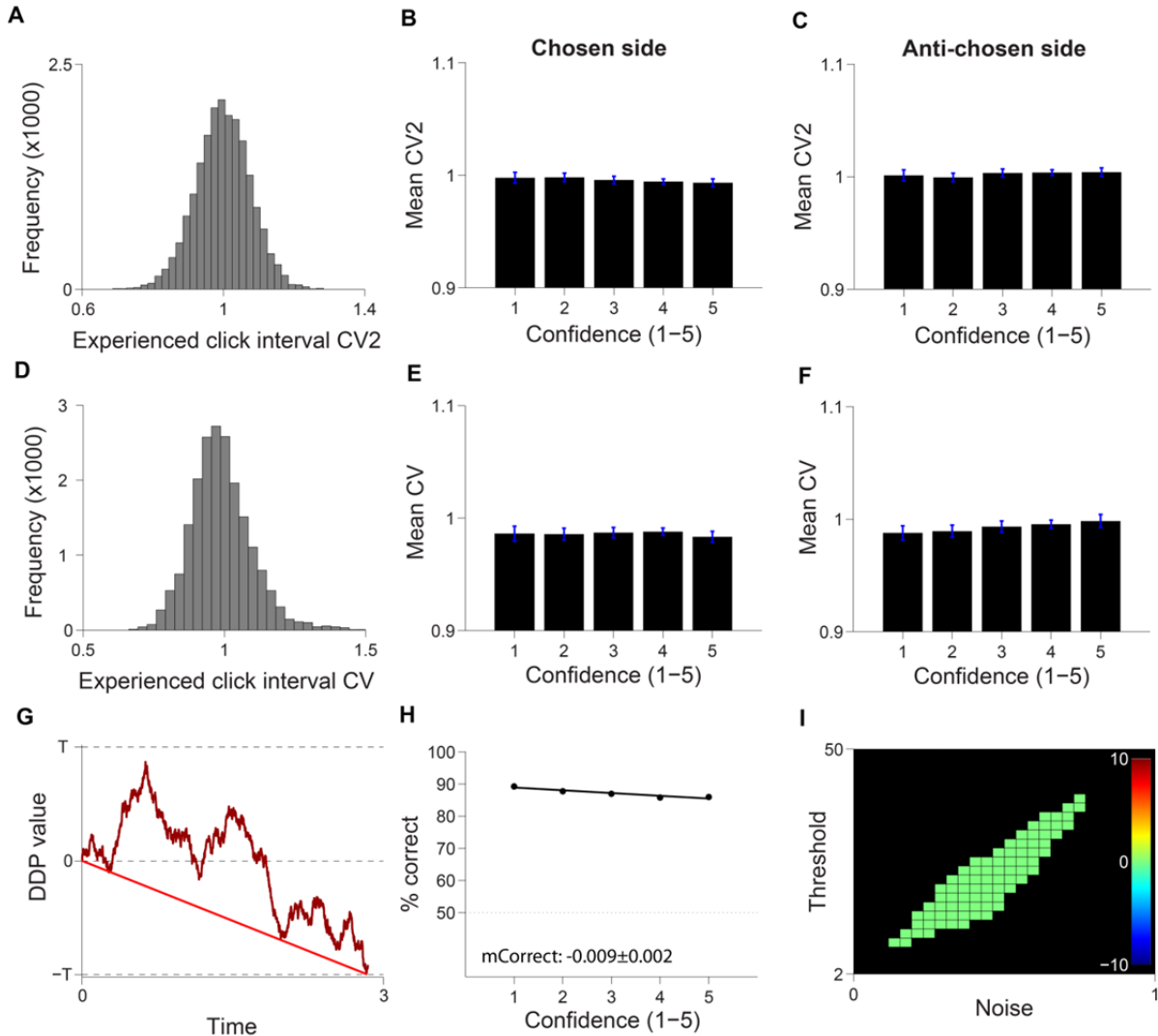


Figure 5.2: Variability of Poisson evidence does not account for decision confidence. For each click train sampled by our subjects, the CV2 and coefficient of variation (CV) were computed for the click intervals experienced, to determine whether confidence is correlated with the reliability of evidence. **A:** The distribution of CV2 for individual trials shows that subjects experienced a range of evidence irregularity. **B,C:** Despite the range on individual trials, mean CV2 was not differentiable among divisions of the confidence reporting scale. Errors show 95% CI of the mean. **D-F:** The same analysis is shown using CV, a simpler measure of variance that is sensitive to changes in event rate. **G:** Illustration of the drift diffusion decision variable time course on a trial with no variance in momentary movement at each dt (bright red) and a trial with high variance (dark red) of the same reaction time. **H:** The inverse standard deviation of momentary movement of the decision variable was evaluated as a confidence measure. For the fitted drift diffusion model shown in section 5.2, this confidence measure did not predict accuracy. Points show model accuracies for each scale division. Trend line shows linear regression fit. Text inset shows slope and bootstrap confidence interval. **I:** Across a range of parameters that produce plausible accuracy and reaction times for our dataset (see section 5.2), inverse variance of decision variable movement was not positively correlated with accuracy. Color bar inset shows slope of the same fit in panel H. Subject values ranged between 7.1-13.3.

We have shown that verbal confidence is insensitive to the range of click interval distributions experienced by our subjects. We have further shown that in the drift diffusion framework, variability in the drift diffusion decision variable's time course does not predict outcomes, failing the most essential test for a confidence measure. We conclude that the characteristics of variance, when considered apart from reaction time and balance of evidence, do not match the patterns of confidence reports in our data.

5.4. Accumulator models predict an incorrect pattern in use of recent evidence

Sequential sampling models of choice (reviewed in section 1.4) posit that decision makers aggregate noisy evidence for competing hypotheses over time until the instant some criterion is reached. The different stopping rules used in these models make different predictions about the patterns in evidence for and against the chosen hypothesis immediately prior to a choice.

We were originally attracted to the accumulator model of choice and confidence (Vickers 1979), for its simple explanation of confidence as the difference between accrued evidence for and against the chosen hypothesis at the moment of choice (dE). The same patterns in normative, statistical and human confidence were also shown to be generated by the dE measure in the race model (Kepecs, Uchida et al. 2008). However, in an accumulator model, an absolute stopping rule terminates accumulation of evidence when either decision variable exceeds threshold.

To determine which patterns in recent evidence are generated by these stopping rules, we computed the reverse correlation of chosen-side and anti-chosen side evidence leading up to the moment of choice for an accumulator model, a drift diffusion model, and an example test subject using the method described in section 2.2.5. We determined a set of each model's parameters that fit every point on a model psychometric to within 95% binomial confidence intervals of the example subject's psychometric function (Figure 5.3a), while the model's reaction time mean was within one standard error of the subject mean. For the model and subjects, we separately aligned all chosen and anti-chosen click trains to the moment of choice on evidence-neutral trials (50Hz left/50Hz right), and

determined the mean click rate in 100ms bins going backwards in time. We subtracted the baseline click rate (50Hz) from chosen and anti-chosen streams in each bin to view selection for elevated or suppressed momentary click rates on each side. The resulting reverse correlation is shown superimposed on the human reverse correlation in Figure 5.3b for the accumulator model, and 5.3e for the drift diffusion model. In reverse correlations produced by both models, evidence diverges strongly from baseline in the first 100ms bin before choice. However, unlike the model, human subjects must execute a motor response following choice. To simulate motor latency, we added two free parameters to each model, describing the mean and variance of a Gaussian latency added to each reaction time, and re-determined parameters that fit our data within error. The resulting reverse correlations are shown in Figure 5.3c,f. While the drift diffusion model provides a fit within error, accumulator model choices select for an incorrect pattern in anti-chosen evidence: click rate on the anti-chosen side is not preferentially suppressed in the moments before choice.

We observed that the absolute stopping rule makes three predictions about the structure of evidence leading to the moment of choice in our task. 1. While the value of the anti-chosen decision variable must have been less than the chosen variable at the moment of choice, its time course leading up to that point is not otherwise filtered against by the stopping rule. For an accumulator model with Gaussian decision variable noise and no decay parameter (perfect memory), this predicts uniform enhancement of chosen-side evidence and uniform suppression of anti-chosen side evidence over time. 2. In a task where subjects sample evidence for a self-determined duration (reaction time task), trials with lower than average absolute evidence in the beginning of the evidence stream take longer to reach threshold. Thus, the earliest evidence in the reverse correlation is expected to fall below baseline. 3. In a regime where noise is low, a final pulse of evidence causes the decision variable to exceed threshold more often than noise. This predicts that in windows measuring the instant before choice, the average click rate of the click stream on the chosen side will be higher than in earlier windows.

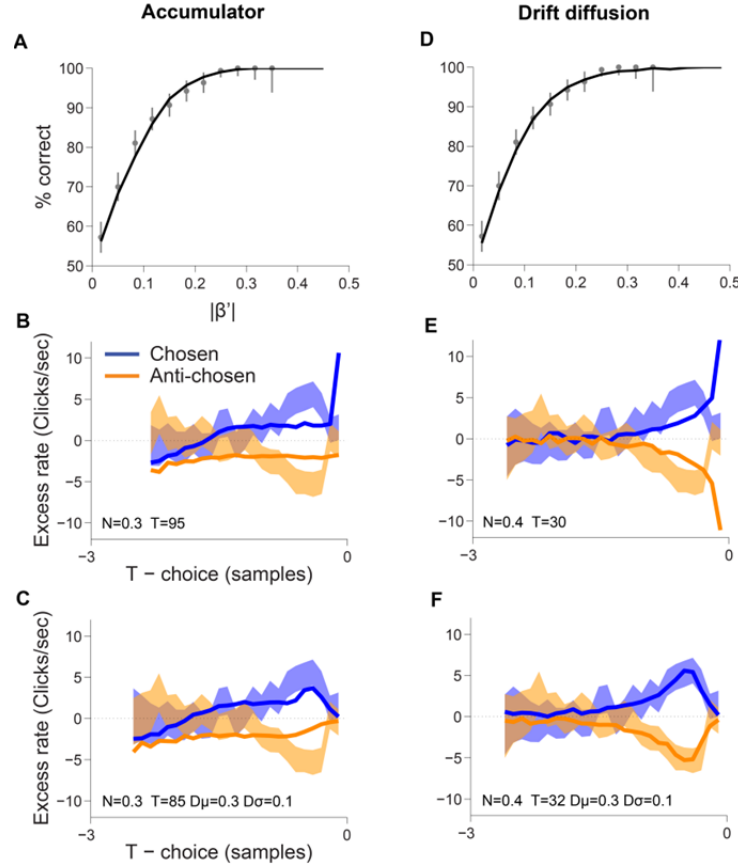


Figure 5.3: The accumulator model incorrectly predicts the reverse correlation of evidence.

A,D: Psychometric fits for two-parameter accumulator and drift diffusion models respectively to an example subject (Subject H1). Free parameters were noise level (N) and threshold (T). Subject data are shown as gray points and 95% binomial CI, solid black lines were produced by model datasets of 100,000 trials. **B,E:** Reverse correlations for the accumulator and drift diffusion models respectively (solid lines), overlaid on the human reverse correlation (shade = 95% CI of mean). Reverse correlations were computed using evidence-neutral trials (50Hz/50Hz) and show the average difference between click rate and 50Hz baseline rate (excess rate) in 100ms bins, when Poisson click streams were aligned to choice at time 0. Text insets show model parameters that fit the human psychometric curve shown to within error of each point, and produced reaction time means within 1 standard error of the human mean. **C,F:** Two-parameter Accumulator and drift diffusion models were extended to four parameters to account for human motor delay by adding a Gaussian motor delay mean ($D\mu$) and variance ($D\sigma$) to each model. The drift diffusion model (Panel F), but not the accumulator model (Panel C) fits the human reverse correlation of evidence that contradicted the hypothesis (Anti-chosen).

To quantify the suppression of anti-chosen side evidence in the moment before choice in each subject, we computed their reverse correlations separately. Chosen-side click rates were significantly enhanced and anti-chosen click rates were significantly suppressed in the moments before choice, indicating that each subject listened for both of these features when making choices (Figure 5.4a-d).

In two subjects, the absolute magnitude of chosen-side enhancement and anti-chosen side suppression were not statistically differentiable (Figure 5.4e) and the imbalance was small but significant in the remaining two subjects.

While the pattern of opposing chosen and anti-chosen side excess rates suggests that a relative stopping rule (or something very close to this) was used by our subjects, the pattern in both our subjects and in the drift diffusion model (Figure 5.3f) becomes increasingly less distinguishable beyond ~300-400ms pre-decision. We wondered whether the predictive information contained in click streams about choice followed the same time course. To determine this, we divided the click streams preceding choice into 150ms bins, and computed receiver operator characteristic (ROC) curves to determine how well left/right click balance in each bin predicted choice using the AUC measure (Figure 5.4f, see methods). Between 0 and 150ms pre-choice, click balance did not inform choice in our example subject (Figure 5.4G) or for the pooled population (Figure 5.4H). The most predictive window was 300-450ms pre-choice, consistent with the greatest divergence between chosen-side enhancement and anti-chosen side suppression. From this window, the information about choice contained in click balance decayed with a comparable time course to chosen/anti-chosen excess rate divergence (Figure 5.4a, g).

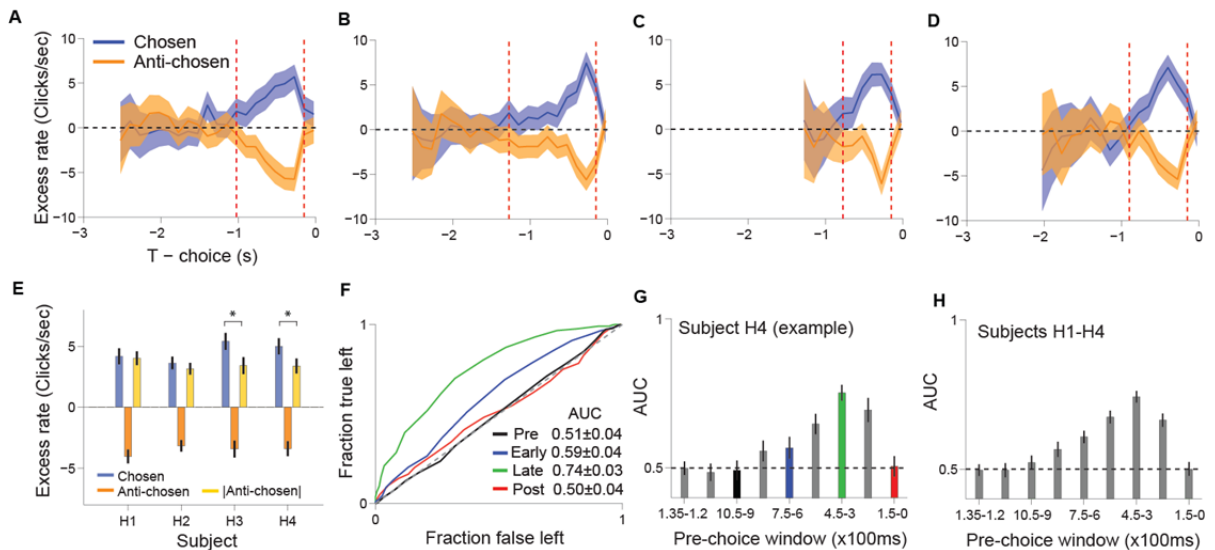


Figure 5.4: The human stopping rule filtered for supporting evidence and against contradictory evidence in the choice-informative final second of sampling. **A-D:** Reverse correlations of Poisson click evidence are shown for subjects H1-H4 respectively, aligned to the detected button press at time 0. On neutral-evidence (50Hz/50Hz) trials, excess click counts on the chosen (blue) and anti-chosen (orange) sides were computed in 50ms bins. The first and last bin which are statistically differentiable at 95% confidence (bootstrap test) are indicated with red dotted lines. This window was used to compute symmetry of aggregated evidence. **E:** Chosen and anti-chosen reverse correlations are nearly symmetric. Excess rate is shown for each subject, averaged throughout the evidence window computed in the previous analysis. Yellow bars are anti-chosen evidence, inverted for visual comparison. Only for subjects H3 and H4, the absolute magnitude of anti-chosen side suppression was less than the magnitude of chosen side enhancement. **F:** ROC curves measure how well click balance predicts choice in an example subject, in four equally spaced 150ms bins. The axes indicate at each click balance threshold, what fraction of left trials was correctly labeled left (ordinate), and what fraction of right trials was mislabeled left (abscissa) by left/right click balance in four windows. The area under each curve (text inset) indicates how well evidence from the partition predicted choices. A classifier that predicts high and low confidence at chance accuracy would have an area indistinguishable from the identity line (gray, dashed). Click balances in the final 150ms before button press and 1 second prior to choice were not predictive, while click balance in intermediate windows predicted choice. **G:** For the same example subject, AUC is shown in each consecutive 150ms window. Windows from the previous analysis are indicated in color. Errors are bootstrap 95% confidence intervals. The predictive power of evidence follows approximately the same time-course as the subject's reverse correlation (panel A). **H:** The same analysis for all subjects pooled. Evidence in the final 150ms before choice was not informative.

We have shown that the accumulator model predicts that recent evidence for the anti-chosen hypothesis is *not* impoverished in the moments before choice – while this pattern is clearly present in our data. In light of this finding, we were forced to abandon the accumulator model, despite its intuitive way of producing confidence reports. We have also shown that the time courses of chosen-side excess rate enhancement and anti-chosen side excess rate suppression match the time course with which evidence balance predicts choice, consistent with the hypothesis that the balance of clicks (and not an absolute threshold) was the primary feature in evidence which our subjects used to drive decisions. In the next section, we extend this analysis to determine whether the time-course of click contribution to *confidence* matches the time-course of contribution to choice, in the context of a model that uses only post-decisional evidence to compute confidence.

5.5. Post-decision models overestimate how well post-decision evidence predicts confidence

Reaction time and variance failed to provide a convincing method of generating confidence from a drift diffusion model. A recent study in humans proposed and tested a novel extension of the

drift diffusion model to account for confidence. In this model, a subsequent phase of evidence gathering is commenced at the moment of choice, for the purpose of determining confidence (Pleskac and Busemeyer, 2010, reviewed in section 1.5.5). According to this model, termed 2-Stage Dynamic Signal Detection (2DSD), *only* post-decisional evidence contributes to confidence. Since our data contain a precise record of the time course of evidence, we wondered whether post-decisional evidence was more informative about confidence than earlier evidence. To determine how much evidence from each click stream was used to compute confidence, we subdivided each 150ms bin from our reverse correlation analysis of choice (Section 5.4) into “high-confidence” (4,5) and “low-confidence” (1,2,3) partitions. We observed that in the final moments before a choice response was registered, high confidence trials select for click balances supporting the same side on average, and low confidence trials select for click balances supporting the *opposing* side (Figure 5.5a). This pattern is consistent with a subject’s use of post-decisional evidence gathered during the motor response in computing confidence. We reasoned that if choice and confidence estimation occur in series, high and low confidence trials would not be distinguishable during the phase of sampling prior to the motor response. To measure the first moment for which high and low confidence are distinguishable in our record of evidence, we combined the chosen and anti-chosen data from Figure 5.5a to generate the time series shown in Figure 5.5b. We then determined the first bin of evidence within the window significant for choice, which also had significantly different high and low confidence means. Evidence contributing to confidence was present in the earliest fifth of the choice window, and persisted until the response was registered.

In estimating the motor response latency of our subjects (section 3.3), I showed that clicks in the final 150ms before the button press response do not inform choice – a finding consistent with the reverse correlations computed in Figure 5.3a-d and our ROC analysis of the same bin in Figure 5.4h. If post-decisional evidence was really used by subjects to compute confidence, we would expect to find that the click balance in this window can be used to predict confidence with greater than chance

accuracy on evidence-neutral trials. To test this, I selected four divisions measuring the pre-choice period, early and late choice periods, and post-decisional period as in Figure 5.4G.

For each trial, I computed the difference between chosen and anti-chosen clicks in each window. To determine whether this quantity had information about confidence, I scored confidence responses as “High” (4,5) or “Low”(1-3) and used an ROC (Receiver Operating Characteristic) analysis (see methods) to determine how well differential click count in each time bin can predict high or low confidence. For an example subject, the ROC curve for each window is shown in Figure 5.5d, with the AUC statistic and its bootstrapped confidence interval printed in the panel inset. While evidence in the pre-choice period could not predict confidence above chance accuracy, evidence in all other windows predicted confidence. The time course of L/R click balance choice prediction and Chosen/anti click balance confidence-prediction AUC statistics are shown in Figure 5.5e-f for two example subjects in sequential 150ms bins. In the pre-choice period (1350-1200ms before choice), neither predictor exceeds chance. While *choice* AUC drops to chance prediction in the final 150ms, this bin predicts confidence for both subjects. Thus, in the post-decision period, evidence *only* informs confidence. The same analysis is shown for our combined human population in 200ms bins in Figure 5.5g. The larger bin size was chosen for statistical power, to illustrate that the early bins informing choice *do* also contain information about confidence.

Computational models of choice and confidence make different predictions about use of post-decisional evidence. Since only post-decisional evidence informs confidence in the 2DSD model, evidence gathered in the beginning of a choice should be agnostic with respect to confidence. This prediction is shown in Figure 5.5h-i, for 2DSD model simulations of 15,834 and 250,000 trials respectively, with parameters loosely fit to our psychometric, chronometric and confidence data. The larger dataset was illustrated in 50ms bins, to illustrate the relative time courses with which evidence informs choice and confidence. The initial bins for which evidence informs confidence occur when the stimulus stops becoming more informative about choice, as the increasing choice commitments begin to divert average information flow from choice to confidence. The post-decision bin, when

virtually all trials no longer inform choice, informs confidence vastly more than evidence collected during the choice period for the 2DSD model (Figure 5.5i). Human post-decision bins in our individual and pooled subjects are comparable to the later bins in the choice period, suggesting a discrepancy in the way confidence is computed from evidence.

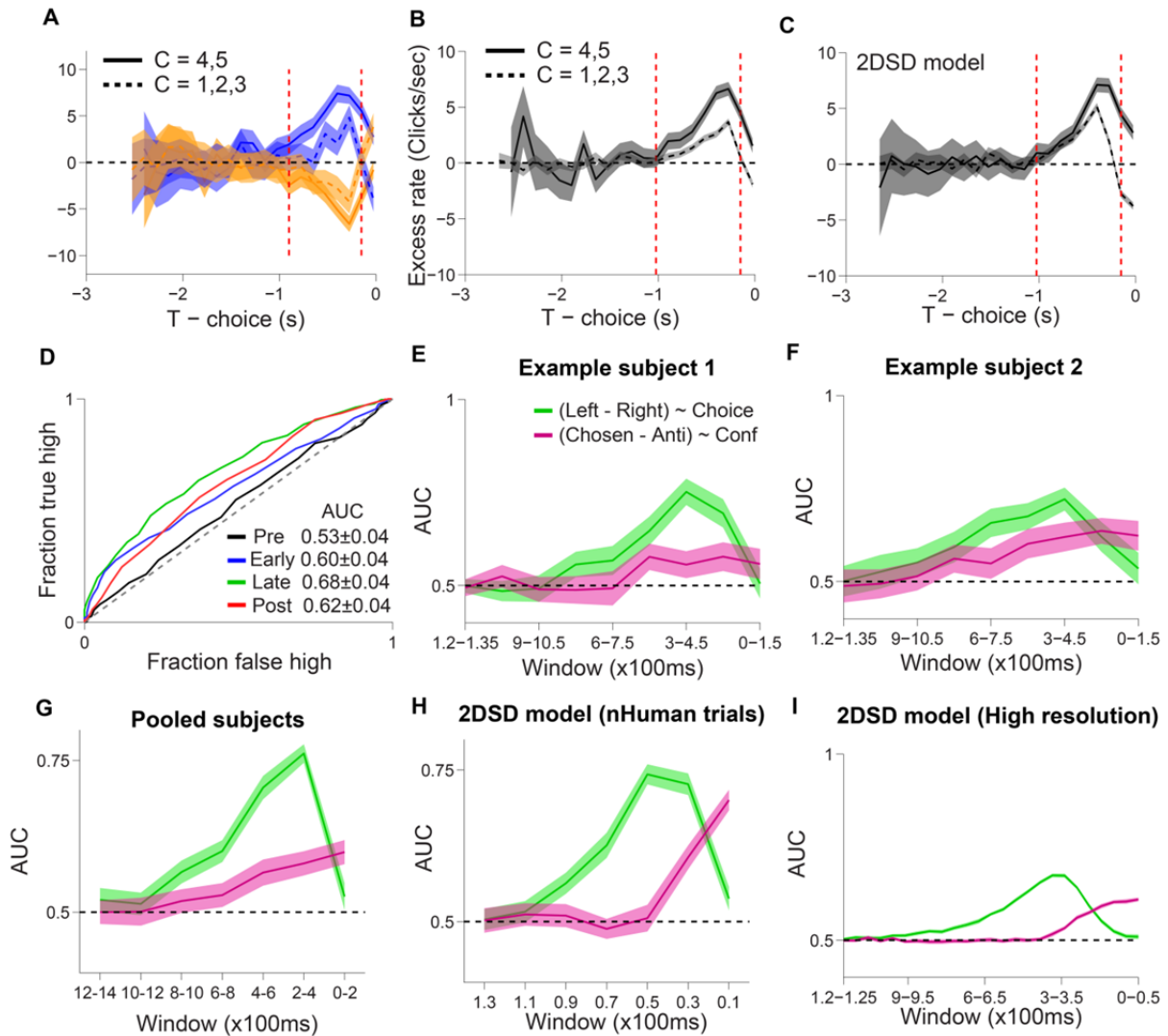


Figure 5.5: Confidence is computed using evidence gathered both during and after choice. A-D: Reverse correlation for the merged population decomposed by high and low confidence. High confidence (4,5) is indicated with solid lines, and low confidence (1-3) is indicated with dotted lines. High and low confidence trials are differentiable throughout the choice process. In the final moments before the motor response, evidence on high confidence trials supports the chosen hypothesis on average, while evidence on low-confidence trials supports the opposing hypothesis. **G:** The analysis from panel F, with high and low confidence components pooled. The first statistically differentiable

bin between high and low confidence is indicated with a blue dotted line. **H:** The four partitions used for ROC analysis, to determine whether evidence from different epochs informs confidence (see main text). The post-decisional evidence bin duration is 150ms, while the other bins are 350ms. **I:** ROC curves for each partition in an example subject. The axes indicate at each threshold, what fraction of high confidence trials were correctly labeled high confidence (ordinate), and what fraction of low confidence trials were mislabeled high (abscissa) by chosen/anti-chosen click balance. The area under each curve (text inset) indicates how well evidence from the partition predicted confidence reports. A classifier that predicts high and low confidence at chance accuracy would have an area indistinguishable from the identity line (indicated in gray). With the exception of the pre-choice partition, all partitions contained information about confidence. **J:** For the same subject, a comparison of confidence predicted by chosen/anti-chosen click difference and choice predicted by left/right click difference is shown. Post-decisional evidence informs confidence but not choice. Early evidence in choice informs confidence. **K:** The same analysis for all subjects pooled. **L:** A confidence-partitioned reverse correlation generated by a 2DSD model simulation with parameters loosely fit to our psychometric, chronometric and confidence data. Only evidence from the part of choice that overlaps with the motor latency distribution contributes to confidence. The first bin distinguishable between high and low confidence is indicated with a black dotted line.

We observed that in humans performing our intended task, as evidence increasingly informed choice, it also informed confidence (but see appendix III for a counter-example with a different decision making strategy). We also observed that post-decisional evidence informed confidence in humans. This finding poses a challenge for the previous models in this chapter that either rely on post-decisional evidence alone, or forfeit the opportunity to exploit post-decision evidence.

5.6 Coupled accumulator models cannot fit both recent evidence and confidence patterns

The accumulator model can be viewed as the end of a continuum of models between race and drift diffusion. Intermediate models are defined by a coupling coefficient, ρ (Figure 1.3), which controls the degree to which evidence added to one decision variable is subtracted from its opponents (Usher and McClelland 2001). Two subsequent adaptations of the coupled accumulator model posit that the balance of evidence measure in these intermediate models provides useful confidence information (Moreno-Bote 2010; Zylberberg, Barttfeld et al. 2012). We found that at $\rho=0$, the model is an accumulator model and is capable of fitting the relationship of confidence to evidence, but not the reverse correlation of anti-chosen evidence (Figure 5.6a-b). At $\rho=1$, the model is a drift diffusion

model. Since the difference between winning and losing accumulators at the moment of choice is fixed by the relative stopping rule, each trial has the same confidence by definition (Figure 5.6g-h). Two intermediate fits are shown, illustrating that as ρ increases, anti-chosen side suppression of evidence increases but the slope of error trial confidence in evidence decreases, eluding a fit to the subject's patterns in evidence and confidence. In section 5.7, we introduce the OCA model - a derivative of the coupled accumulator model that uses evidence collected after commitment to a motor response to fit human confidence patterns despite high coupling coefficients.

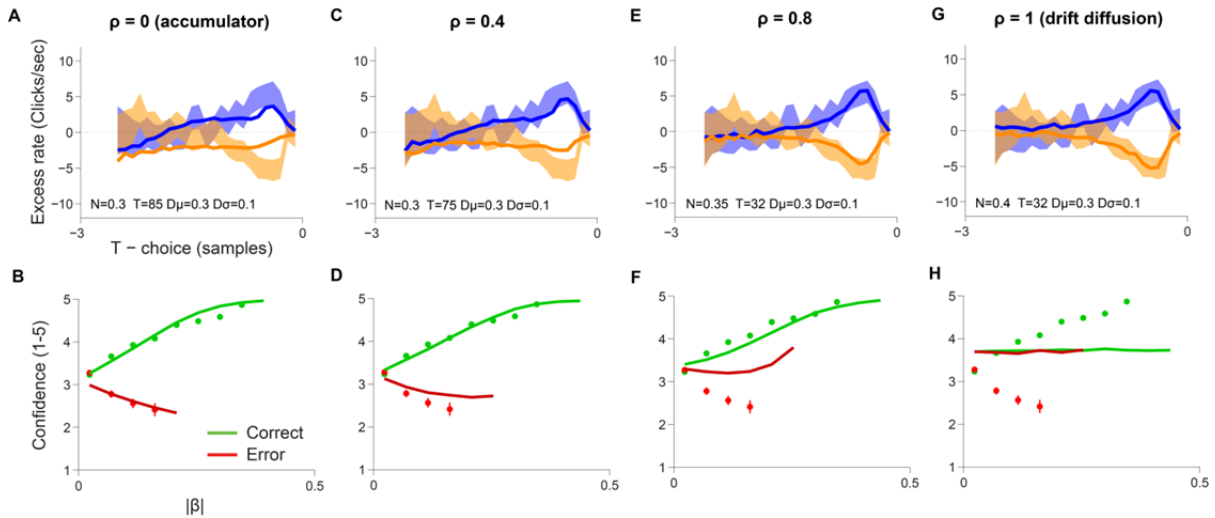


Figure 5.6: The coupled accumulator model cannot simultaneously fit patterns in evidence history and confidence. In the top row of panels, the reverse correlations of chosen and anti-chosen evidence are shown for subject H1 (blue = chosen side, orange = anti-chosen side, shaded area = 95% CI of mean, insets show parameters of model fit) and for 150,000-trial simulations of the coupled accumulator model evaluated at coupling coefficients equal to accumulator, two intermediates and drift diffusion with remaining parameters fit to psychometric and reaction time data as in Figure 5.1. The bottom row shows confidence v. discriminability at each coupling coefficient (points = subject mean, errors = 95% CI, solid lines = model). **A,B:** The accumulator model provides a poor fit to the reverse correlation of evidence (also shown in Figure 5.3), but provides a reasonable fit to decision confidence v. evidence on error trials. **C,D:** At $\rho=0.35$ (mostly accumulator), slight selection for a low click rate in anti-chosen evidence is apparent, but the slope of error trial confidence v. discriminability becomes too shallow to fit our data. **E,F:** At $\rho=0.8$ (mostly drift diffusion), the anti-chosen side reverse correlation of the model approaches error of our data, but the slope of error trial confidence v. evidence becomes slightly *positive*. **G,H:** The drift diffusion model fits the reverse correlation of evidence (also shown in Figure 5.3) but each trial is terminated with chosen and anti-chosen accumulators at the same threshold, producing the same confidence on each trial.

Aside from difficulty simultaneously fitting these two clear patterns, the coupled accumulator model has a fundamental problem it shares with the reaction time, variance and accumulator models – these models are not capable of gathering post-decisional evidence to inform confidence, while it is clear that human decision makers do so. In Figure 5.7, I show the time course of confidence and choice-predictive evidence generated by human decision makers and each of the models considered until this point when fit within error to the psychometric function and reaction time. Only the 2DSD model correctly predicts that post-decisional evidence informs confidence.

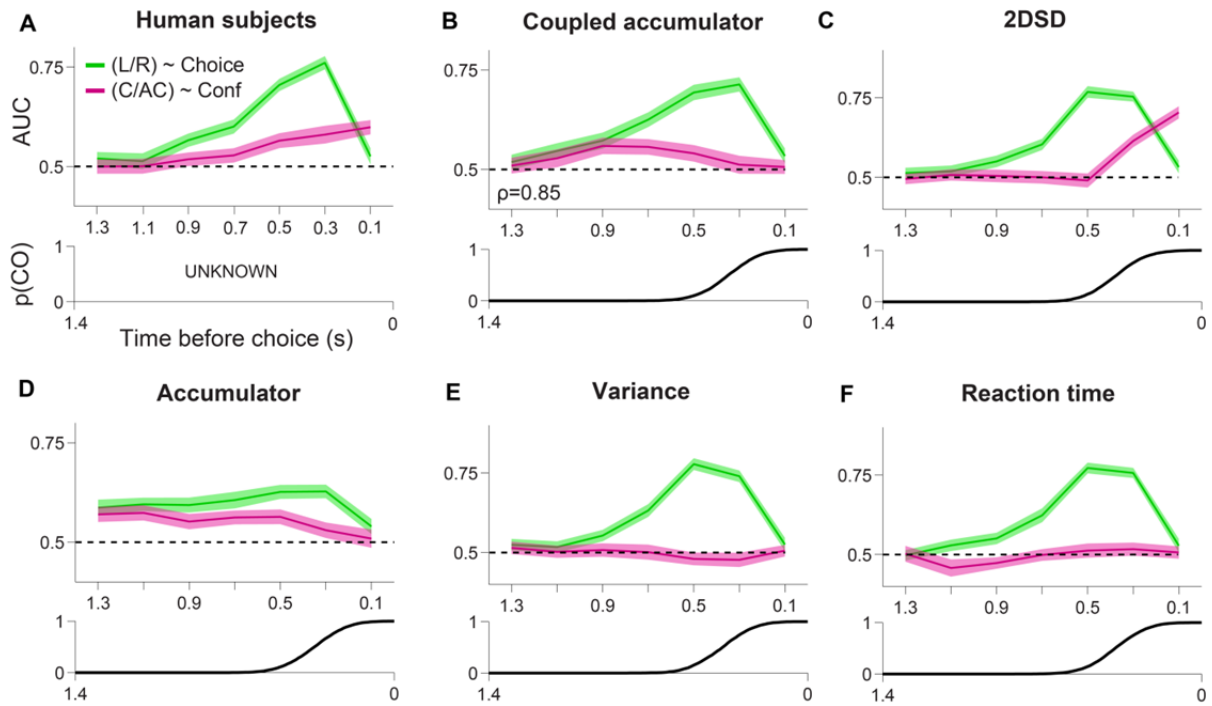


Figure 5.7: No previous model correctly predicts temporal patterns of evidence contribution to choice and confidence. AUC was computed to determine how well left/right click balance predicted left/right choice (green), and how well chosen/anti-chosen click balance predicted high(4,5) vs. low (1-3) confidence trials (purple) in 200ms time bins preceding choice. All trials used were evidence-neutral (50Hz/50Hz). Abscissa labels indicate 200ms bin centers in seconds. Shade indicates bootstrap 95% CI of mean AUC (solid lines). The lower plot in each panel shows the fraction of choices that have committed to a motor response for models: the cumulative sum of the commitment/latency distribution. To approximate the error of our human data, model data were generated from simulations the same size as our human dataset (15,834 trials). **A:** In human subjects, evidence increasingly informs confidence from at least the first 1/5th of the period that informed choice until the button press (also see Figure 5.5). The distribution of commitment times is unknown. **B:** The coupled accumulator model was shown for $\rho=0.85$, a value that could fit the reverse correlation of evidence while also generating some confidence information (see Figure 5.6).

Confidence information was gathered from the earliest moments of choice, but returns to AUC=0.5 with increased motor response commitment. **C:** The 2DSD model uses only post-commitment evidence, producing a pattern where evidence does not predict confidence in early choice, and is too informative in the final 200ms bin where nearly all choices are committed. **D-F:** The race, variance and reaction time models return to chance prediction of confidence with increased motor response commitment.

We noticed that the confidence pattern produced by human subjects (panel A) appeared provocatively similar to a merge of unique Coupled Accumulator and 2DSD model properties (panels B and C respectively): escalating use of evidence in the early period that informed choice, and use of post-decisional evidence.

5.7 Our OCA model simultaneously fits interrelationships of five key behavior measures

Encouraged by the partial successes of previous models to account for select features of human choice behavior, we developed a new model. It contains features of two previous models that were each insufficient: the Coupled Accumulator model and the Two-stage Dynamic Signal Detection (2DSD) model. We term this model the Opportunistic Coupled Accumulator (OCA) model. The model's behavior for collection of evidence occurs in two phases: 1. integration of evidence until commitment to a decision, governed by equation set 5.1 (adapted from equation 1.5 in Chapter 1), and 2. less efficient integration of evidence during the commitment and motor response period, governed by equation set 5.2. The equations for these phases differ only in the stopping rule that terminates the phase, and efficacy of sensory evidence. The interplay between the model's six free parameters and equations is illustrated in Figure 5.8. The coupled accumulator model equations describing the choice phase of evidence collection are:

$$\begin{aligned} L_{i+1} &= L_i + E_{L,i} + \eta_{L,i} - \rho(E_{R,i} + \eta_{R,i}); \\ R_{i+1} &= R_i + E_{R,i} + \eta_{R,i} - \rho(E_{L,i} + \eta_{L,i}); \\ \max(L_i, R_i) &< T \end{aligned} \tag{5.1}$$

where L and R are decision variables, E_L and E_R are vectors of Poisson evidence for left and right hypotheses at each time point, η_L and η_R are noise vectors for the left and right decision variables respectively, drawn independently at each time step i from a standard Gaussian distribution with free parameter N describing variance for choice-period noise. Two additional free parameters are: T (the choice threshold in units of evidence) and ρ (the coupling coefficient, see Figure 5.8a). The choice is determined by the larger of decision variables L , R when the inequality $\max(L_i, R_i) < T$ is satisfied.

To compute evidence collection during the motor response period, the decision variables retain their values from the moment of choice (L_D , R_D) and proceed to integrate more evidence until the end of a Gaussian-derived response latency interval tL :

$$\begin{aligned}
L_C &= L_D + \sum_{i=1}^{tL} S_P E_{L,i} + \eta_{L,i} - \rho(S_P E_{R,i} + \eta_{R,i}) \\
R_C &= R_D + \sum_{i=1}^{tL} S_P E_{R,i} + \eta_{R,i} - \rho(S_P E_{L,i} + \eta_{L,i}) \\
tL &= \mathcal{N}(\mu, \sigma)
\end{aligned} \tag{5.2}$$

This formulation preserves variable names from equation set 5.1 with additional variable tL (the post-decision motor response latency), and adds the latter three free parameters: μ (the mean of the Gaussian distribution of motor response latencies), σ (the variance of the same distribution) and S_P , a factor describing the reduced signal strength of evidence gathered in the post-decision and motor response period, as it is apparent to the accumulators. \mathcal{N} is a function that produces Gaussian numbers with parameters μ, σ . At the end of latency ($i = tL$), decision confidence is determined by the absolute difference between decision variables L and R at the end of the post-decision interval (L_C , R_C):

$$\begin{aligned}
\text{if } L_D > R_D, \quad C &= (L_C - R_C) \\
\text{if } L_D < R_D, \quad C &= (R_C - L_C)
\end{aligned} \tag{5.3}$$

An example of the time course of decision variables for a single trial is shown in Figure 5.8b, with the decision variables at time-points L_C/R_C and L_D/R_D indicated.

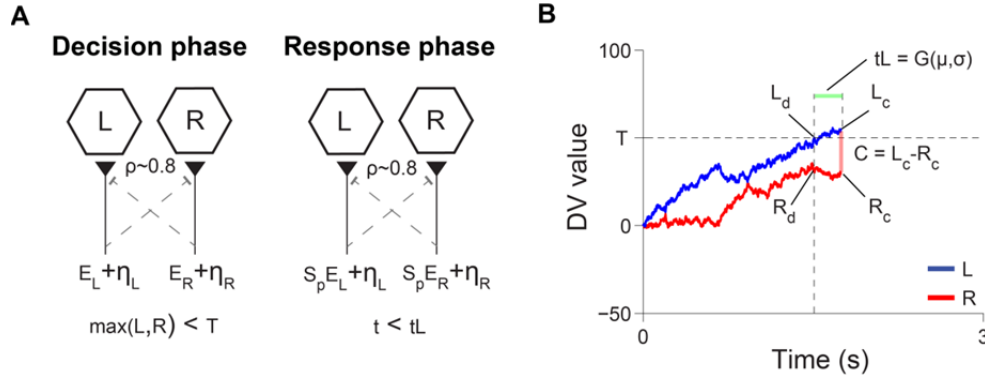


Figure 5.8: Illustration of OCA model variables and parameters in an example trial. A: The relationship between decision variables is shown for the decision and response phases of evidence collection. Decision variables L and R accumulate evidence for respective hypotheses that the left and right directions have stronger evidence. At each time step i , left and right decision variable noise for each sample, η_i is added to any available evidence at each time step, E . For each unit of noisy evidence added to L or R , the same amount multiplied by coupling coefficient ρ is subtracted from the opposing decision variable. The decision phase ends when L or R exceeds threshold T . The response phase is identical to the decision phase except that evidence efficacy factor S_p diminishes the evidence added to each accumulator at dt , and the phase ends when time elapsed exceeds response latency t_L . **B:** The time course of decision variable values is shown for an example trial. Variables were computed using equation 5.1 until one variable exceeded threshold, indicated by the intersection of the left decision variable (blue) with the horizontal dashed threshold line). On threshold crossing, a random response latency t_L (green bar) was drawn from a Gaussian distribution with parameters μ, σ . Subsequent decision variable values were computed throughout the latency period with equation 5.2. Confidence was computed following the latency period as the difference between the chosen and anti-chosen decision variables following the response phase, shown as a light red bar.

To test the model's predictions against the combined data of our test subjects, we adjusted the six free parameters until the model produced patterns describing the interrelationships of choice, confidence, discriminability and the evidence time-course as close as possible to our data. This fit is shown “globally” in 10 projections for a model stimulation of 100,000-trials in Figure 8.9. To produce the model dataset for the fit, click streams used by the model were matched in absolute click rate and rate balance frequency to those in our pooled human dataset. To produce decisions and confidence reports, the model was evaluated once per millisecond, and proceeded to integrate evidence as per equations 5.1 and 5.2. Raw confidence values produced by the model were mapped to

the 5 point scale based on frequency of human scale usage. Since subjects had considerably variable reaction time distribution shapes and means (Figure 3.4), their combined reaction time distribution was unusually shaped. For this reason, reaction times were fit to subjects individually. Three additional projections showing simultaneous fits to reaction time interrelationships are shown with individual subjects' global fits in Figures 8.10-8.14. In nearly every instance, the OCA model simultaneously produced *and fit* each of the patterns in our data which had eluded previous models.

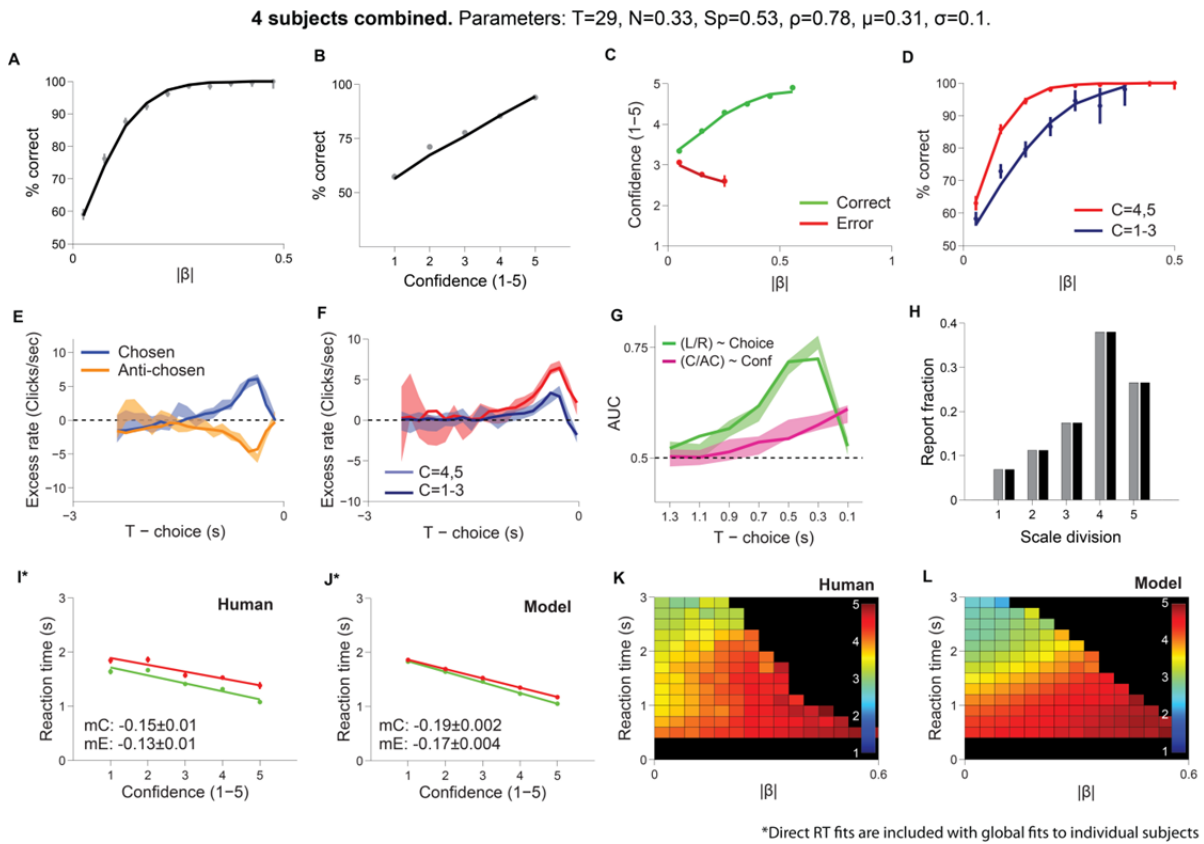


Figure 5.9: The OCA model simultaneously fits patterns in human behavior that eluded previous models. A 100,000 trial dataset was produced by an OCA model simulation, which generated choices and confidence reports from Poisson click stream pairs matching pooled subject data. Patterns from the model dataset were overlaid on patterns in pooled data from our 4 subjects. **A:** The model's psychometric function predicted the subject's accuracies in each of 10 equally spaced evidence strength bins, to within 95% confidence intervals of human data. Solid line shows model accuracies. Gray points and errors show human accuracy and 95% binomial CI. **B:** OCA model confidence strongly predicts accuracy, and fits four of five confidence scale divisions to within 95% binomial confidence intervals of human accuracy. The excessively accurate second bin in pooled human data was attributable to subject H4 (Figure 5.10). **C:** OCA model confidence shows the normative model prediction that the lowest confidence occurs on strong, high evidence trials and

highest confidence on correct, high evidence trials. It further fits human data for the same trend to within 95% confidence intervals of mean confidence in each of six accuracy bins, with the exception of a slight deviation beyond error in a single point. **D:** The model generates the third pattern derived from statistical confidence, that confidence predicts accuracy for a given evidence strength, and fits the human pattern of the same trend. **E:** On evidence-neutral trials (50Hz/50Hz) the OCA model's choice-aligned reverse correlation of evidence showed elevated click rate on the chosen side, and suppressed click rate on the anti-chosen side, decaying in both directions from peak excess rate difference at ~350ms to baseline (50Hz) at approximately 150ms and 1.5 seconds prior to choice. This trend fit human data in 150ms bins (error shade indicating human 95% CI of mean excess rate) with the exception of a single bin on the anti-chosen side, which showed a sharp peak. **F:** When separated by confidence and the anti-chosen stream's suppression inverted and recombined with chosen data to show the amount of excess rate supporting the chosen hypothesis in different time windows, evidence immediately before choice supported the chosen hypothesis on high confidence trials, and contradicted the chosen hypothesis on low confidence trials, consistent with use of post-decision evidence. The model pattern matched human data in this projection at every point in 150ms bins. **G:** The area under ROC curves calculated for each bin to determine how well left and right click balance predicted choice on individual trials was plotted in green. Shaded region shows 95% bootstrap confidence intervals of the AUC for human subjects. Solid line shows AUC for the model dataset. A separate AUC computed to determine how well chosen/anti-chosen click balance predicted confidence in the same bins is plotted in pink for the subject (shade) and the model (solid line). For both the subject and the model, evidence gathered throughout the choice contributed to confidence. **H:** The frequency of confidence scale division use is shown for the subject and the model. This equivalence was fixed by our method of mapping of raw confidence values produced by the model to the 5-point reporting scale. **I-J:** Confidence was significantly anti-correlated with reaction time for both the subject and model. Since subjects had markedly different reaction time distributions, fits interrelating reaction time with other variables are shown for individual subject global fits (Figures 5.10-5.14). **K-L:** Heat maps showing the interrelationship of discriminability, reaction time and confidence are shown for the subject and model respectively. For both the subject and model, confidence varies with both reaction time and evidence.

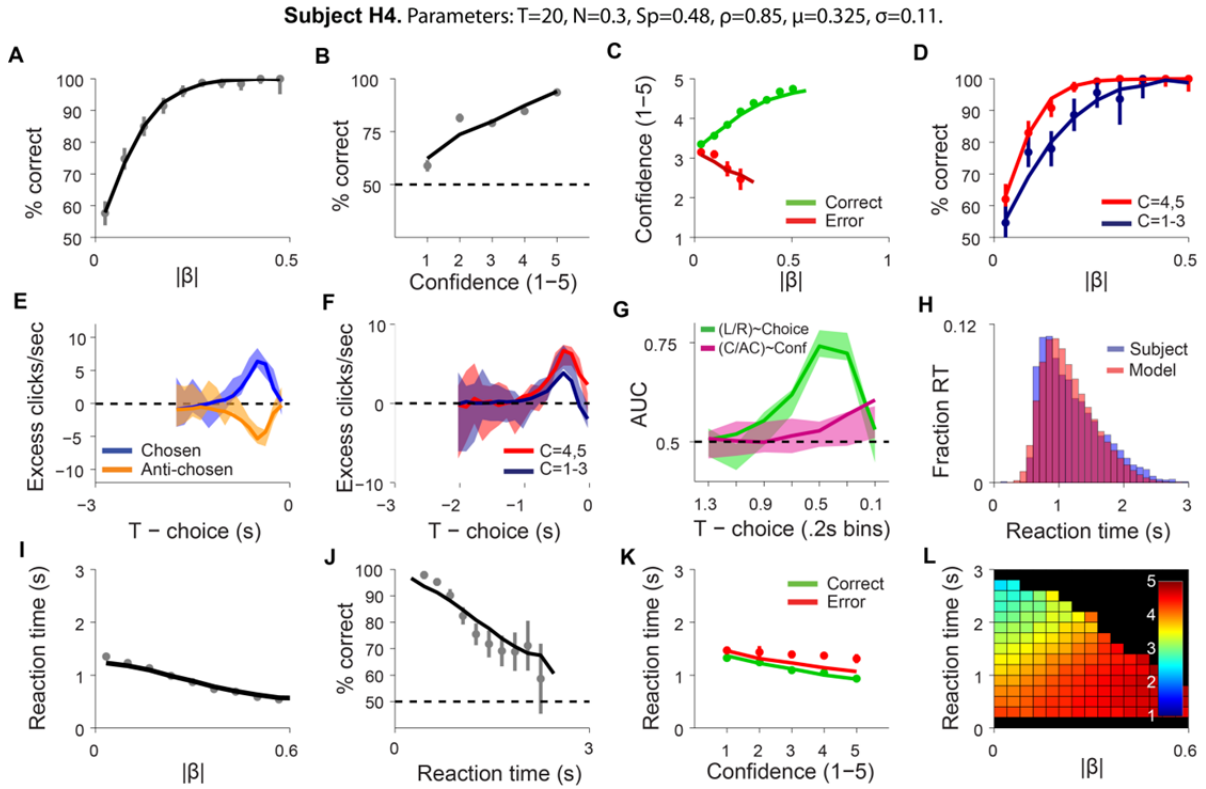


Figure 5.10: For individual subject H4, the OCA model simultaneously approximates 12 key interrelationships between decision measures. A 100,000 trial stimulation of the OCA model using parameters indicated in text above the figure panels, was overlaid on patterns produced by subject H4 as in Figure 5.9. Panels **A-G, L** show the same patterns as in Figure 5.9. **H:** The subject reaction time distribution is shown in blue, and the model RT distribution in red. Mean subject RT was 41ms slower than model RT: $1.098 \pm 0.015s$ and $1.058 \pm 0.002s$ respectively where error indicates 95% CI of the mean. **I:** OCA model reaction time was anti-correlated with discriminability, and matched subject data to within 95% confidence interval of the mean in 8 of 9 evenly spaced discriminability bins. **J:** OCA model choice accuracy decreased as a function of reaction time, though was accuracy was significantly different from the accuracy of human data in 4 of 10 evenly spaced reaction time bins. Error shows binomial 95% CI. **K:** Confidence is shown as a function of reaction time as in Figure 5.9i-j, with subject and model overlaid. OCA model reaction time decreases as a function of confidence, and matches human data to within 95% CI of the mean on correct trials. While the model underestimates the reaction times of error trials, error trial reaction times are longer than for correct trials for trials with the highest confidence – a divergent trend that is also apparent for subject H4. **L:** The interrelationship between reaction time, discriminability and confidence is shown for the model. This projection was previously shown for subject H4 in Figure 5.1f, showing the simultaneous correlation between confidence and discriminability, and anti-correlation between confidence and reaction time.

Subject H3. Parameters: $T=17$, $N=0.25$, $Sp=0.35$, $\rho=0.75$, $\mu=0.35$, $\sigma=0.08$.

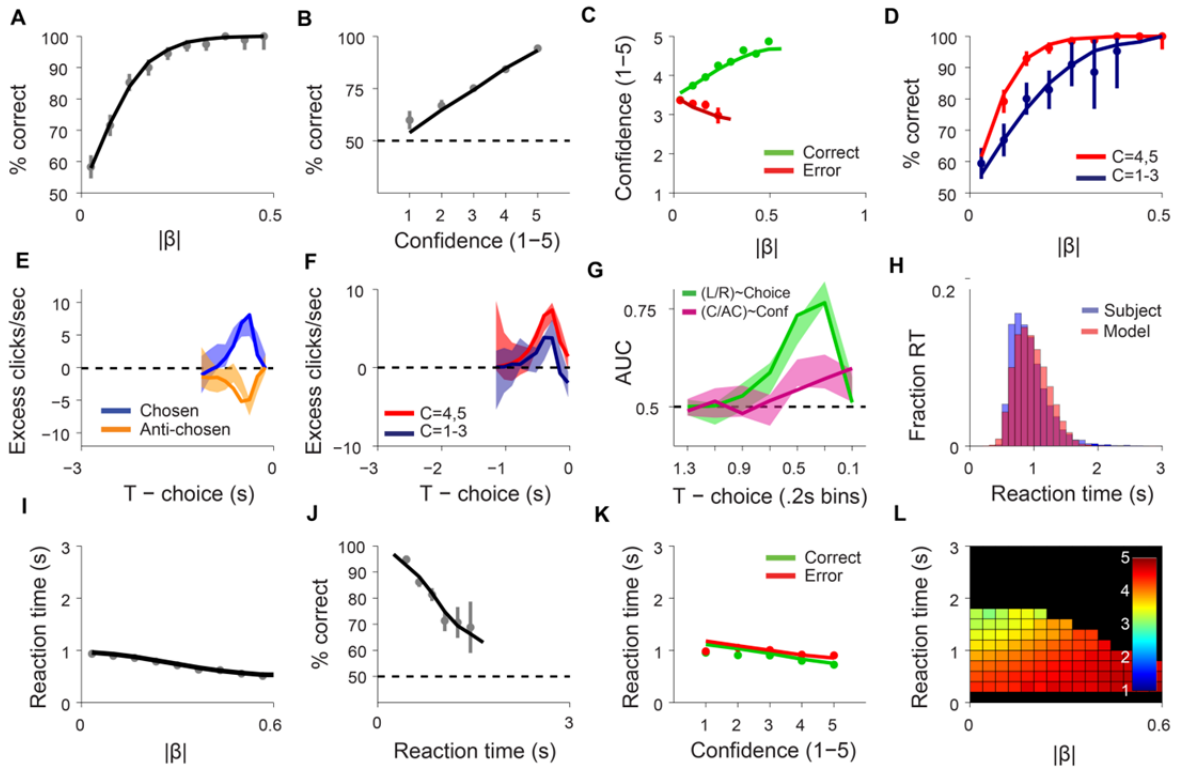


Figure 5.11: OCA model fit for subject H3

Subject H2. Parameters: $T=32$, $N=0.32$, $Sp=0.85$, $\rho=0.82$, $\mu=0.28$, $\sigma=0.08$.

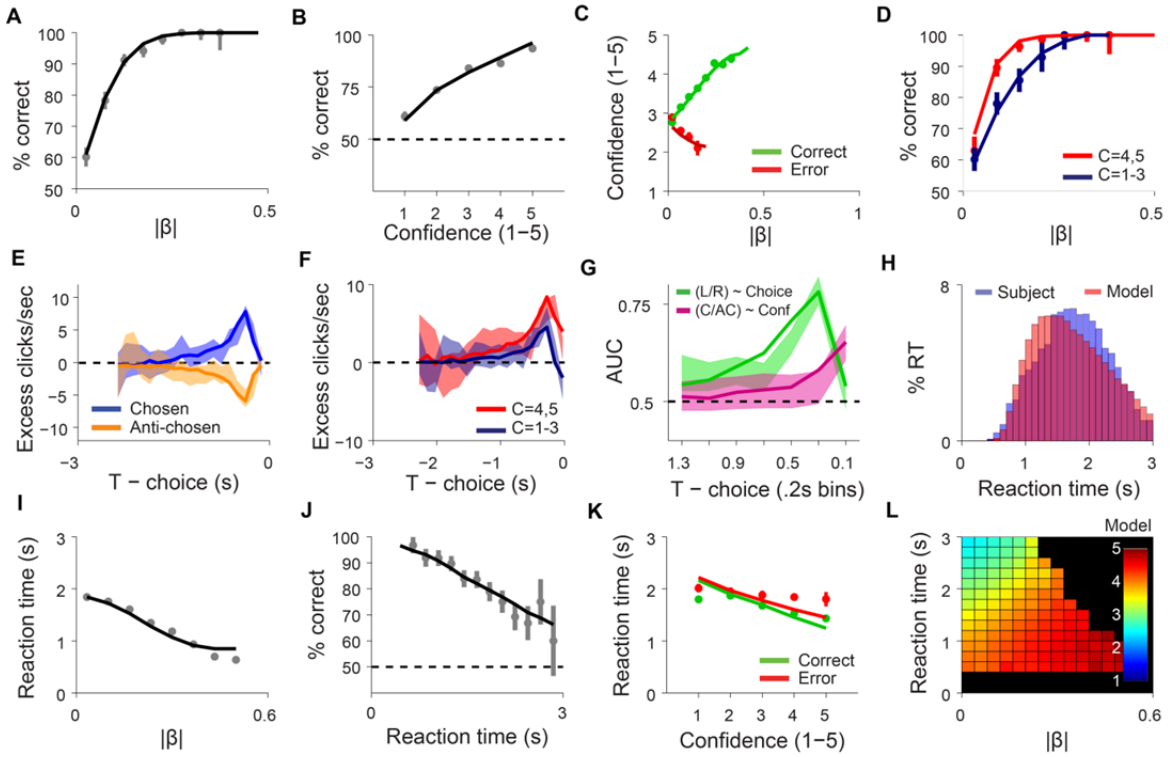


Figure 5.12: OCA model fit for subject H2

Unlike the previous two subjects shown, Subject H2 sampled for long durations on average, and produced a nearly symmetric reaction time distribution which our model had difficulty producing (Figure 5.12H). This is consistent with the possibility that as the subject approached the three second time limit and the probability of a time-out reward forfeiture increased, the subject made decisions in urgency with less evidence supporting their choice. While 7% of trials had reaction times greater than 2500ms, a very small fraction of trials that actually timed out – less than 1%, compared to 8.6% for the model. For our final subject, Subject H1, 15% of all trials approached the timeout boundary, exceeding 2500ms in sampling time while only 1.2% of trials timed out. To achieve a fit of the timeout-approaching reaction time distribution in subject H1 without a large fraction of model timeouts, we added an additional parameter to the model: λ_T , a decay factor for the choice threshold. Once each millisecond, the initial threshold (55.0) was multiplied by $\lambda_T = 0.99965$ to generate a simple, exponentially collapsing decision threshold which was equal to 19.2 at the three second timeout. This enabled our model to generate long reaction times within 79ms of subject H1's reaction time mean, while timing out only 0.2% of trials (Figure 5.13h), and preserving a fit to normative model confidence patterns, psychometric and chronometric functions within error (Figure 5.13a-d, j). Suggestions of how to further improve reaction time fits in situations with subjective hazard are provided in section 5.8.

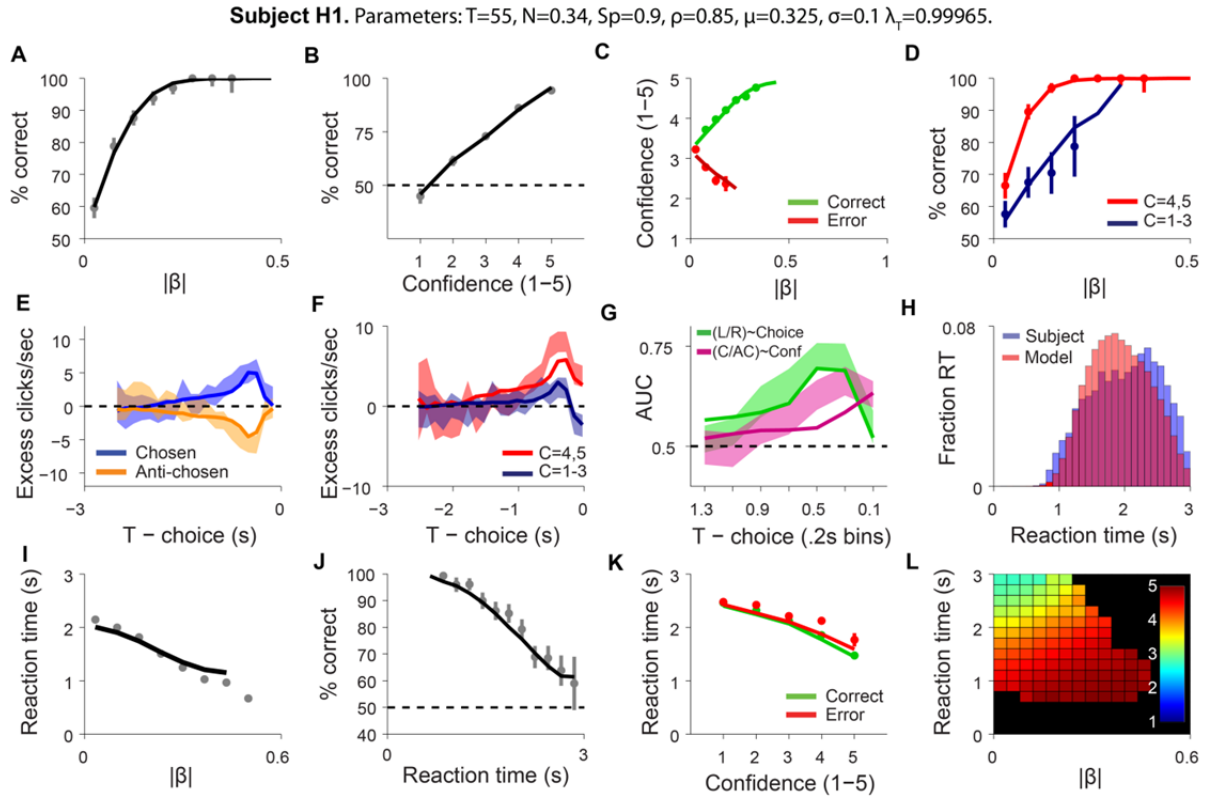


Figure 5.13: OCA model fit for subject H1. To model subject H1, initial decision threshold T decayed by a factor λ_T when the model was evaluated each millisecond. This addition to our model, prevented a high rate of time-outs, though did not provide ideal fits to the history of evidence (panels E-G).

5.8 Conclusions

We began our research with a confidence reporting task that is uniquely capable of accounting for patterns in the history of evidence use at high resolution, and an affinity for the accumulator model, for its straightforward confidence provision and relevance to the physiology of decision making in monkeys (Mazurek, Roitman et al. 2003; Churchland, Kiani et al. 2008). We were initially surprised when it failed to fit basic patterns in evidence history. Subsequently, we discovered that each of the five published models introduced in this chapter (as well as several preliminary models of our own) failed to explain at least one clear pattern in our data – a situation we did not expect to face, given the large literary priors on confidence and decision making. The model we eventually converged upon simultaneously accounts for the interrelationships of choice outcome,

confidence, discriminability, evidence history and to a lesser extent, reaction time – a kind of “global” fit that has been painfully absent in the heated, century-old literature of confidence modeling.

Our Opportunistic Coupled Accumulator model is *opportunistic* in the sense that it can harvest confidence information from the same presented evidence that informed choice, and is separately equipped to gather additional, noisy information during the motor response if this opportunity is available. This dual-source design makes the algorithm useful across a wider range of decision making contingencies than previous algorithms – it is capable of determining confidence in circumstances where the decision maker’s choice policy uses a relative stopping rule to choose, and also in weak signal circumstances where the evidence gathering penalty elicited by a motor response makes further signal collection impossible. This opportunity-driven use of evidence is consistent with the brain’s ability to flexibly use available evidence to optimize learning and performance (Courville, Daw et al. 2006).

OCA is a novel “meta-algorithm” describing how the brain computes choice and confidence. However, the principles governing separate stages of the algorithm are not unique. The balance of evidence collected by decision variables in coupled accumulators has previously been suggested to inform confidence (Moreno-Bote 2010; Zylberberg, Barttfeld et al. 2012), though neither study addressed the trade-off between confidence information and the coupling coefficient (Figure 5.6), or tried to fit confidence reports produced by this model to human confidence. Post-decision evidence collection in a perfectly coupled drift diffusion model had also been suggested to account for decision confidence (Pleskac and Busemeyer 2010), though the model’s over-dependence on post-decision evidence and under-dependence on early evidence with respect to human evidence use patterns (Figure 5.5) makes the conjecture that the brain uses only post-decision evidence unlikely.

OCA especially excels at modeling the relationships between choice outcome, confidence, discriminability and evidence history (Figure 5.9). However, we noticed that the reaction time distributions produced deviate slightly from those produced by individual subjects, subtly affecting fits inter-relating reaction time with other measures. We suspect that our three-second limit for

evidence sampling may have affected the duration for which subjects elected to continue sampling evidence, particularly those subjects who sampled for long periods and frequently approached this limit. A more sophisticated model of the interplay between the cost of gathering evidence and reaction time was proposed by Drugowitsch, Morento-Bote et al, 2012, providing a way to combine both the cost of gathering evidence and the subjective hazard for the cost and momentary risk of “timing out” a trial into a cost function that is then used to determine when to stop sampling for a given evidence stream. We chose *not* to build these computationally demanding features into our model, to demonstrate how surprisingly well our simple algorithm of six-parameters could explain behavior.

Given its surprisingly unique ability among sequential sampling models to explain our high-resolution choice and confidence data, we anticipate that our OCA model will be exploited to provide an algorithmic understanding of momentary brain function underlying choice and investment behavior in a broad range of tasks.

Chapter 6

A device to capture multiple choice decision responses in head-fixed mice

In section 1.6, I introduced the challenge of porting the 2AFC paradigm to the head-fixed setting. This would allow researchers to leverage the latest optical and molecular tools to determine the neural correlates of perception, decision making and motor control, using an informative, precisely controlled behavior. In this chapter, I introduce a novel way for head-fixed mice to communicate their decisions: the choice ball - a modified computer trackball that rotates about an axis, and is operated by short lateral motions of the animal's front paws. I demonstrate the advantages of this response device using an auditory discrimination task, and show that head-fixed mice can be trained in a robust two alternative choice paradigm yielding psychophysical measures.

6.1. Design of the choice ball device and experimental apparatus

The choice ball system employs a series of modifications to a commercially available USB trackball (Expert Mouse model K64325, Kensington). The 55mm diameter trackball is large enough that a laboratory mouse can position both paws comfortably and stably on its surface. We replaced the original trackball with an oversized 55mm diameter ping-pong ball (Joola, USA) in order to reduce the force required for the mouse to rotate it. In order to restrict the trackball's motion to rotation about a single axis, we secured a precision-aligned hypodermic tube through the ball, creating mechanical guidance to the two acceptable choice responses. The hypodermic tube was anchored to metal inserts mounted in grooves cut into the trackball chassis such that the ball could freely rotate about this axis and remain within range of the trackball's optical sensor. To prevent undesired visual cues, we disabled the trackball's time-delayed automatic LED shut-off feature by clipping the trackball LED's leads from its printed circuit board, re-wiring it to an external power source, and securing it back on to the circuit board with epoxy. The raw trackball position was read out by a dedicated PC running the Windows XP operating system (Microsoft) using a script in MATLAB r2008a (Mathworks).

Pointer enhancement was disabled in the operating system to ensure a linear readout of the trackball position. In our setup, a single pixel registered by the trackball computer corresponded to 0.26 degrees of rotation, or 0.13mm of lateral movement along the circumference. Ball position was monitored and logged to the computer's hard drive in real-time. During each trial, the computer measured ball movement towards pre-determined choice boundaries. Choice boundary crossing events were communicated in real-time from the trackball computer to our behavior system using parallel port logic lines. Position data were returned over an Ethernet connection following each trial, to the computer governing our behavior system. The modified trackball module itself was thus inexpensive and easy to integrate with our existing behavior system, based on BControl (C. Brody, Z.F. Mainen, A.M. Zador, CSHL) , an open source real-time state machine framework (Felsen and Mainen 2008; Erlich, Bialek et al. 2011). The Matlab code used on the dedicated trackball computer to acquire trackball position data and interface choice responses with our behavior system is provided in a GitHub repository:

<https://github.com/KepecsLab/ChoiceBallSystem>. On this site we also provide code and setup instructions for a microcontroller-based Choice Ball system that is functionally identical to the one used here but does not require an additional, dedicated computer.

A stable platform to support the hind paws of a mouse was positioned such that when head-fixed, the animal's front paws rested on a modified computer trackball as shown in Figure 6.1a. The platform was made from a polycarbonate tube (McMaster, 50.8mm outer diameter) cut to 100mm in length and milled to expose the animal's front paws to the ball as depicted in Figure 6.1b (a photo diagram of the setup). A custom designed optical lickometer was placed within reach of the animal's tongue. A lickometer is a device specialized for precise delivery of liquid droplets and measurement of lick events. Water reward was dispensed from a reservoir above the animal into the lickometer, through silicone elastic tubing. Water flow was controlled using a solenoid pinch valve (NRResearch, Inc.), with pulse timing calibrated such that a single reward measured 5 μ l. Sounds were delivered using a set of speakers (Harman Kardon 5187-2105) positioned on either side of the animal's head.

Speakers were calibrated to 70dB-SPL within a 5-40kHz range using a pressure-field microphone (Brüel & Kjær, Sound & Vibration Measurement A/S, Nærum, Denmark). Auditory stimuli generated in MATLAB were sampled at 200kHz and delivered to the speakers using a Lynx L22 sound card (Lynx Studio Technology, Inc.). A white LED to indicate trial onset was mounted on the top face of the lickometer, pointing towards the mouse. An infrared camera was positioned 20cm in front of the animal, and its output was shown on an LCD display for monitoring purposes. The trackball was placed on a lab jack (Fisher Scientific) to allow fine scale adjustment of its height relative to the mouse's paws. The entire apparatus was enclosed in a dark acoustic isolation chamber (Industrial Acoustics Company, Inc.).

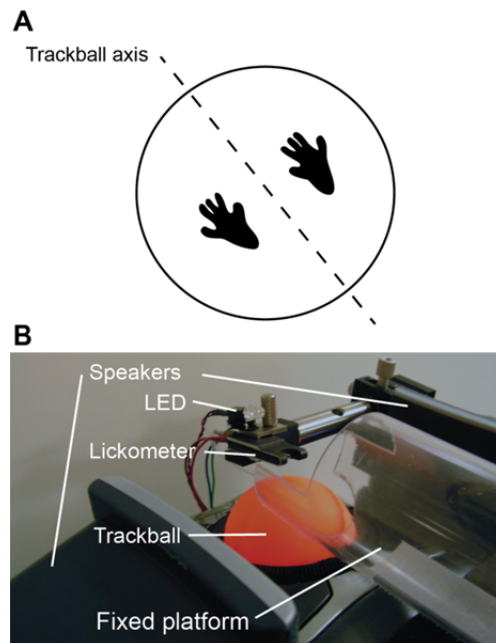


Figure 6.1: The choice ball apparatus. **A:** Orientation of the choice ball. A ping pong ball replacing a commercial trackball was fitted with a steel rod, and secured in the trackball chassis such that it rotated freely about its axis in range of the trackball's optical motion sensor. At the beginning of each trial, the mouse's front paws rested on the ball approximately as shown (also see supplementary video). **B:** Photo diagram of task setup. A head-fixed mouse is positioned with its front paws on the choice ball and its hind paws on a fixed platform. A water-dispensing lickometer is positioned within reach of its tongue. An LED on the lickometer indicated the no-movement period to the mouse at the start of each trial. Speakers positioned laterally on either side of the mouse were used to present the stimulus (head-fixure beams omitted for clarity).

6.2. Design of the task and training protocol

To demonstrate the utility of the Choice Ball device, we trained mice in a two choice Poisson click stream discrimination task similar to our rat and human tasks. We chose this task because it provided graded decision difficulty and directional cues to match directional responses for easy learning. Finally, we sought a task that allowed for hundreds of trials, permitting within-session analysis at each difficulty level. In this section, we describe our adaptation of the Poisson click discrimination task for head-fixed mice.

6.2.1 Subjects and surgery

Data are reported from three male B6129SF2/J white-bellied agouti mice, aged 12 weeks at the onset of training. Mice were housed individually on a 12 hour reversed light/dark schedule. Four additional mice were used but never advanced beyond the earliest phases of training due to low performance in the auditory task and persistent side bias. Food was available *ad libitum*, and the mice were placed on a liquid restriction schedule with daily body weight monitoring to ensure that body mass remained within 85% of mass prior to restriction. Mice were provided with at least 1ml of water per day. One hour following training, if the water delivered in the task did not exceed the daily water allowance, the remaining portion was provided to each mouse in its home cage.

For surgery, mice were anaesthetized with intra-peritoneal injection of ketamine (150mg/kg) and xylazine (12mg/kg). The skull was exposed, and a horizontally oriented titanium bar (20mm x 3mm x 1mm) was centered 2mm above bregma and secured to the skull surface with acrylic cement. Three bone screws spaced among the occipital and parietal bones were used to ensure the stability of the implant, and the remaining exposed skull was sealed with acrylic cement. For post-operative analgesia, ketoprofen (5 mg/kg) was administered intra-peritoneally. Mice were given one week to recover from surgery, with water and food available *ad libitum*.

6.2.2. Task flow of events

Each trial was initialized by illuminating an LED positioned in front of the animal (Fig. 6.2c), to indicate to the animal that they must cease moving the ball for a period of one second in order to continue the trial. This ensured that any paw movements captured during a trial were initiated from resting position, at some time following stimulus onset. While we did not store a record of trackball movement before trial start, we observed in video records that mice frequently shifted their paws, often in the opposite direction of their paw movement in the previous trial as would be necessary to regain balance on the ball. Once a one second period had passed with no ball motion, the LED was extinguished and two Poisson-distributed click trains of different rates were delivered from right and left speakers. Mice listened to independent random click trains, and were rewarded for rolling the trackball towards the side with the faster underlying click rate. Mice were not required to sample the stimulus for a fixed period, and were permitted to respond as soon as they had made a choice. To qualify as a response, mice had to move the choice ball by 26mm ($\sim 50^\circ$) about its axis. Failure to respond within the 2 second stimulus delivery period resulted in termination of the current trial and initialization of the next trial after a three second delay. Mice were rewarded for pushing the ball in the direction of the click train whose underlying rate was faster as shown in Figure 6.2b. A correct response was immediately rewarded with 5ul of water dispensed from the lickometer. An incorrect response resulted in a 5 second punishment delay before the next trial was initialized.

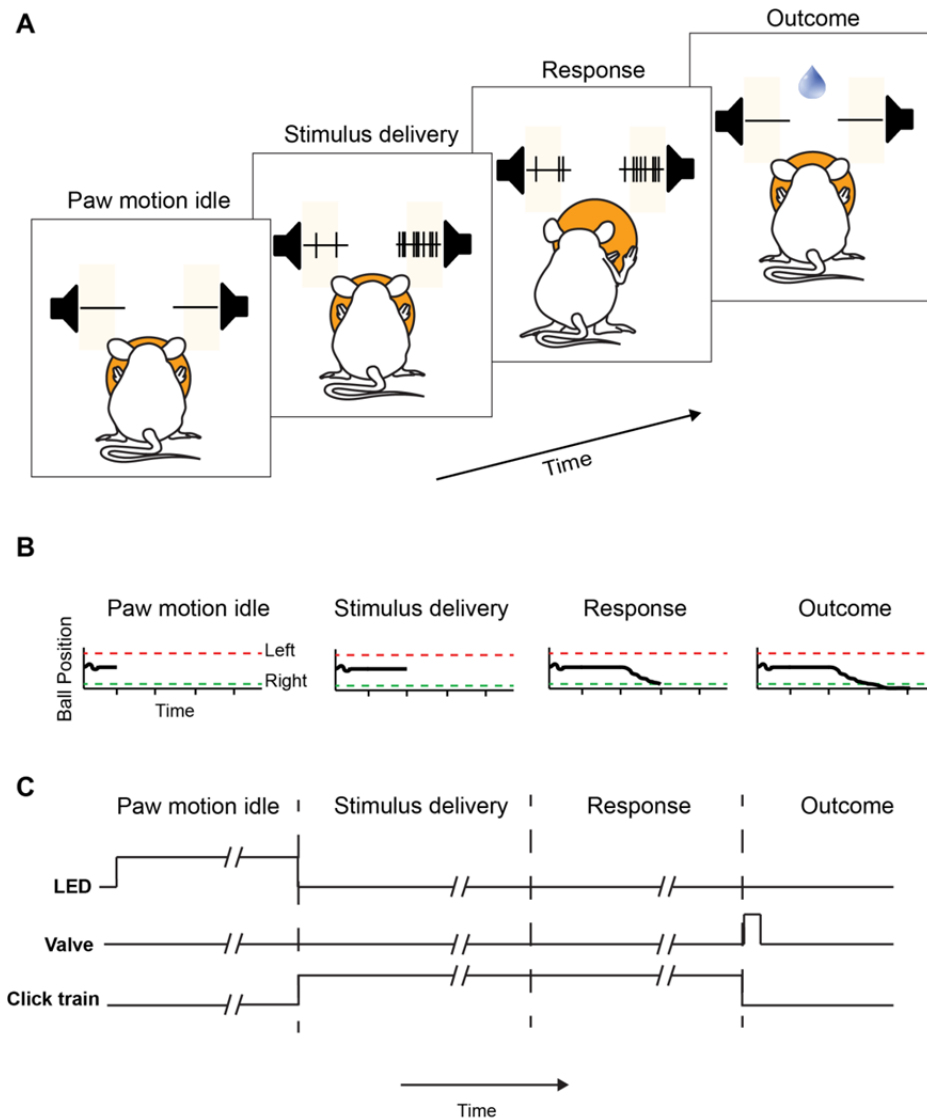


Figure 6.2: A two-alternative forced choice task for head-fixed mice. **A:** Trial stages. 1. Paw motion idle: a head-fixed mouse must cease paw motion for one second to initiate a trial. 2. Stimulus delivery: The mouse listens to two random click streams and discriminates the faster clicking side. 3. Response: The mouse rotates the trackball with its paws in the chosen direction until a response threshold is crossed. The click train was not terminated until a choice was registered, allowing ongoing clicks to inform paw movement. 4. Outcome: The mouse was rewarded with a water droplet if it responded correctly, or punished with a 5 second time-out delay if it responded incorrectly. **B:** Time course of ball position record. In each trial, a fixed threshold was set at 26mm of lateral paw movement ($\sim 50^\circ$ about the ball axis) such that ball rotation past threshold in the correct direction will register a correct response. **C:** Illustrated time course of trial-start LED, water valve and the auditory stimulus.

6.2.3. Training protocol

After surgery recovery, mice were introduced to the apparatus following three days on a liquid restriction schedule. For each mouse, we optimized the orientation of the lickometer prior to each session, such that licks interrupted the lickometer's photo-gate, and the mouse's tongue touched the drink tube when extended. We also adjusted the elevation of the hind-paw platform for each mouse, such that the animal's anterior/posterior axis was parallel to the floor with its front paws resting on the trackball. The objective of training was for mice to use the choice ball device to classify two Poisson click streams by indicating which stream had a faster underlying rate. Training was implemented in six phases. The first phase was designed to train head-fixed mice to lick for water. Mice were trained to lick by intermittently rewarding their licking behavior. Mice were advanced to the subsequent stage once they were rewarded for licking at least 200 times in a single 30 minute session. In the second phase, mice learned to lick preferentially in the period following an auditory cue. Each trial started with a random delay between one and three seconds in length. Following the delay, a one second long train of random clicks was presented at 100Hz mean click frequency from both left and right speakers. Clicks were 1ms white noise pulses, flanked by 100 μ s linear ramps between silence and 70dB SPL. Mice were rewarded for licking in the second following the click-train offset. Mice were advanced once they licked in the second following sound offset for at least 90% of 200 consecutive trials. The third phase was intended to teach mice to use the ball to dispense reward. Parameters were identical to the second phase, except that the click train was extended to two seconds in length, and mice were required to move the trackball in either direction while the click train was being presented, in order to terminate the click train and dispense reward. Sessions in this phase of training were repeated twice daily until mice responded with a 6.5mm lateral trackball movement ($\sim 13^\circ$ about the ball axis) at the correct time for at least 50% of the first 100 trials in a session. In the fourth phase of training, sound was presented from only one side on each trial, and mice learned to push the ball selectively, towards the side with sound. The rewarded side with sound

was alternated trial-wise between left and right. Choosing the wrong direction resulted in repeating the same stimulus on the subsequent trial, until the animal chose correctly (to prevent choice bias). During this phase, we trained mice to move the ball for increasing distances in order to register a response. Mice were initially required to move the ball 6.5mm ($\sim 13^\circ$) in the chosen direction to respond, and once accuracy on this phase surpassed 90% per session, the distance to choice threshold was gradually increased to 26mm ($\sim 50^\circ$, 200px). We initially chose 26mm as the rewarded choice boundary for our task based on early observations that mice were unwilling to paw repeatedly for distances longer than 26mm over hundreds of trials per session (data not shown). A session would end once the mouse failed to cross a choice threshold for ten sequential trials following trial 100. Training on the fourth phase was repeated twice daily until choice accuracy surpassed 90% for a single session of >100 trials with the choice movement boundary set at 26mm. In the fifth training phase, left and right trials were randomly interleaved so that mice could not perform well by simply alternating choices trial-wise and ignoring the stimulus. The anti-bias algorithm was disabled for this phase and all subsequent phases. The sixth training phase (multiple-difficulty) was identical to the fifth, except that Poisson distributed click train pairs were generated at three levels of stimulus difficulty. The rates used to generate the left and right click trains always summed to 200Hz as follows for each difficulty level: 200Hz/0Hz, 0Hz/200Hz (easy left, easy right), 150Hz/50Hz, 50Hz/150Hz (mid left, mid right) and 125Hz/75Hz, 75Hz/125Hz (difficult left, difficult right). Mice were returned to the fifth stage of training (single difficulty) for the remainder of the experiment, following several multiple difficulty sessions. In these subsequent sessions, the boundary for reward delivery was extended beyond 26mm in a manner dependent on performance, and LEDs positioned on top of each speaker indicated the chosen direction to the mouse once the ball passed 26mm. Outcome scoring for these sessions in our analyses used the earlier 26mm decision threshold. We report on three mice that completed training in the 2AFC task. Following the first three phases of training, we collected 68 behavior sessions from our animals, averaging 283 completed trials per session. Although the task is not explicitly self-initiated, mice responded reliably, rotating the

trackball by a sufficient amount to be counted as a valid response (26mm, 200px) on >90% of all trials in behavioral sessions.

6.2.4. Readout of choice responses: mice respond with rapid (~200ms) paw motions

In a decision making study, an ideal choice response is rapid (i.e. a saccadic eye movement), and can be aligned to the record of neural data with enough precision to be informative about brain function at moment of choice. The paw movement component of choice responses was swift, averaging 189ms of ball motion to register a choice (Figure 6.3a,b).

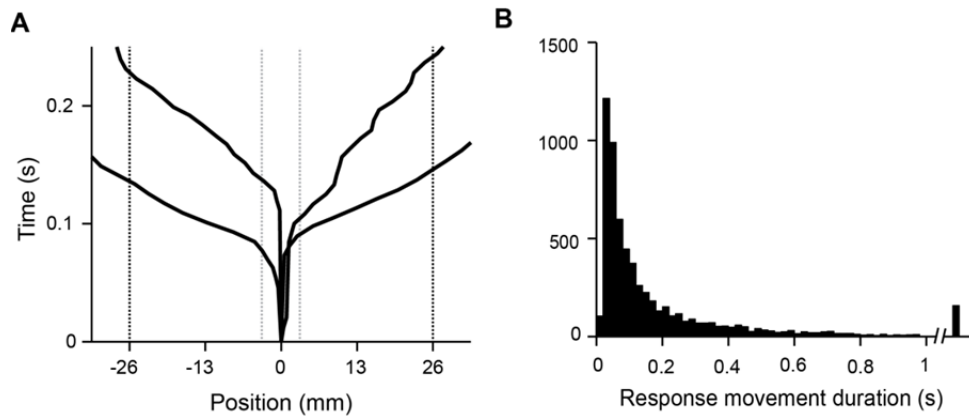


Figure 6.3: Rapid readout of choice responses. A: Trackball position data. The trackball position record allows for high-resolution reconstruction of the animal's motor response. Ball movement trajectories for four example trials are shown with respect to auditory stimulus onset. Choices are registered when the trackball has rotated by 50 degrees in either direction (shown as 26 mm along the ball's circumference with respect to the position at trial start). B: Timing of paw movement responses. Variability in ball position was calculated for the first 50ms following auditory stimulus onset of 10,575 post-training trials of three mice. The standard deviation of position during this pre-reaction period was 1.65mm. The duration of paw sweeps was considered during the period from 2xSD (3.3mm, gray dotted lines in 3a) to the \pm 26mm decision boundary (black dotted lines), as a measure of the speed with which the mice operated the ball to report their decisions. Trials in which a mouse moved 6.5mm in one direction and ultimately responded in the opposite direction were considered decision reversal trials (Fig. 6) and omitted from this analysis. Mean response movement duration was 189 ms.

6.3. Mice perform the task and show evidence of speed/accuracy tradeoff

In this section we report successful training of three mice, showing that the relation of reaction time to accuracy and discriminability is consistent with reports of these relationships in monkey and human psychometric discrimination.

6.3.1. Mice learn the task within weeks

Mice typically learned to discriminate among the easiest stimuli in less than two weeks. Figure 5.4a shows an example plot illustrating the improvement in discrimination accuracy over the course of training for a single mouse. Three types of sessions are shown, reflecting the order and difficulty of trial types: sessions for which trials with the easiest stimuli to classify were alternated trial-wise between left and right correct responses (black), sessions for which the easiest stimuli were randomly interleaved (dark gray) and randomly interleaved sessions containing stimuli of multiple difficulty categories (light gray). Mice typically approached 90% proficiency at classifying the easiest stimulus over the course of the first 5,000 trials as shown in Figure 6.4b. For sessions with multiple difficulties, difficult stimuli were created by configuring the relative rates of clicks coming from the two speaker channels.

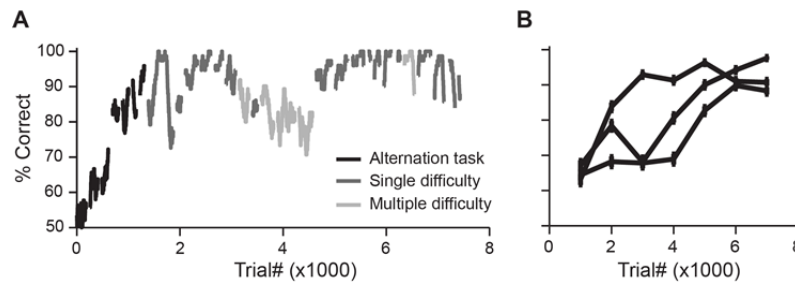


Figure 6.4: Mice learn the trackball task consistently and with high accuracy. A: Sliding window performance for a single mouse (window length = 100 trials). Mice were initially trained, using only the easiest stimuli (100% left / right contrast), to alternate left and right responses (black bins). When accuracy approached 90%, mice were advanced to sessions with interleaved 100% contrast stimuli (dark gray bins), and subsequently sessions with multiple difficulty levels (light gray bins). **B:**

Performance records for three mice, considering only trials with 100% left/right contrast (bin size = 1000 trials). Error bars show S.E.M.

6.3.2. Mice show psychometric discrimination

Psychometric performance varied with discriminability, indicating that mice had mastered the intended click train discrimination task. Data are shown for all 10 multiple difficulty sessions and for a single example session of 466 trials in Figure 6.5a-b.

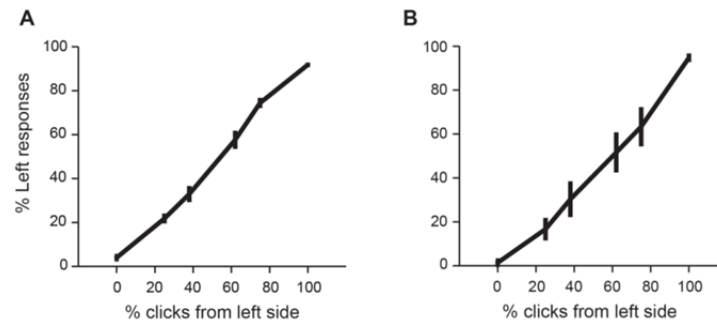


Figure 6.5: Choice accuracy varies with discriminability. A: Pooled psychometric performance for all sessions across mice for sessions with multiple difficulties, showing a full range of accuracy in classifying stimuli of varying difficulty. Left-side click percentage reflects a total of 200Hz distributed between the left and right channels. Error bars show 95% binomial confidence interval. **B:** The bipolar psychometric function for an example session of 466 trials.

6.3.3. Mice sample difficult stimuli for longer, and are more accurate for longer RT

In our task, due to temporal variability in the stimulus, optimal performance requires an evidence integration strategy (Palmer, Huk et al. 2005). Specifically, mice could gain an accuracy advantage by using a decision making strategy where more ambiguous evidence is intentionally sampled for longer periods. With respect to easy stimuli (200Hz contrast), mice spent significantly more time sampling medium (100Hz contrast) and difficult stimuli (60Hz contrast), as shown in Figure 6.6a for all sessions averaged across mice (ANOVA and post-hoc Tukey-Kramer test, $p < 1e-6$). Using the same test, the difference between medium and difficult stimulus sampling time was not significant ($p > 0.05$). For the example session shown in figure 6.6b, the mouse also spent significantly less time sampling easy stimuli than difficult stimuli ($p < 0.05$). An additional prediction

of integration in our task is higher performance on trials where mice spent more time sampling the binaural click train prior to responding. Figures 6.6c-d show that over the first 200ms, accuracy improves as reaction time increases for the easiest stimulus condition across sessions (note non-overlapping 95% binomial confidence intervals for the increasing trend over the first 200ms in Fig. 6.6c and a similar trend for the first 200ms of sampling in the example session in Fig. 6.6d).

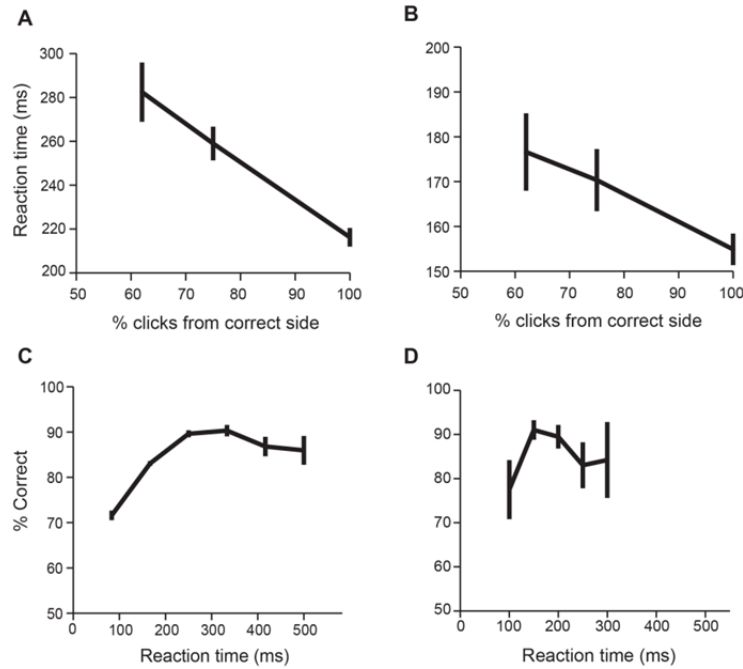


Figure 6.6: Reaction times correlate with stimulus difficulty and performance. **A:** Averaged reaction time with respect to discriminability considering only correct trials across all multiple difficulty sessions, showing that mice spend more time sampling the stimulus before responding on difficult trials. Error bars on plots C-D show S.E.M. **B:** Example session: reaction time with respect to discriminability. **C:** Pooled reaction time conditional accuracy for three animals. Accuracy peaks when animals respond 250ms following stimulus onset. Error bars show 95% binomial confidence interval. **D:** Example session: reaction time conditional accuracy.

6.4. Mice reverse their choices for a performance gain

Beyond precise control over behaviorally relevant stimuli and measurement of discrete lick events, the choice ball affords us an additional window into the decision making process - the opportunity to reconstruct the time course of the animal's report of its decision. In analyzing ball

movement data, we noticed that mice occasionally reversed the decisions they had initiated. In this section, we provide the first evidence of decision reversals in mice.

6.4.1. Readout of choice reversals

If reversal trials were used strategically to correct errors made as part of the decision making process, we anticipated that they would be correlated improved performance, and would become less necessary as the mice became more proficient in the task. To test these predictions, we defined a choice reversal as a response in which the mouse moved the trackball 6.5mm in one direction, but ultimately chose the other direction. The boundary for choice reversals was placed at 6.5mm, corresponding to four standard deviations from the mean of lateral ball movement during the first 50ms following stimulus onset. The time courses of trackball movement for five choice reversal trials are shown in red in Figure 6.7, overlaid upon 45 surrounding non-reversal trials from the same session in blue. In total, 6.9% of all trials qualified as reversals by these criteria.

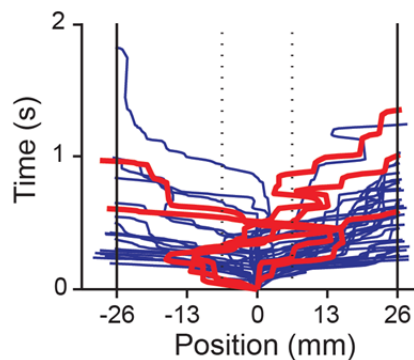


Figure 6.7: Detection of choice reversals in mice. Ball movement time course plot for 50 consecutive trials. Dotted line at ± 6.5 mm indicates the reversal detection threshold. Solid lines at ± 26 mm indicate decision threshold. Detected reversal trials are shown in red.

6.4.2. Mice reverse their choices to optimize performance

Performance on reversal trials within the same difficulty category was slightly but significantly enhanced with respect to non-reversal trials (Fig. 6.8a, Chi-square test, $p < 0.001$), consistent with the prediction that mice can reverse their choices to improve accuracy. If mice use

decision reversals to resolve errors made as part of learning, we reasoned that reversal probability would diminish as a function of experience in the task. For this analysis, we excluded alternation sessions, and considered only sessions with randomly interleaved trials, where the rewarded choice boundary was set at 26mm. The latter restriction was necessary because the decision to reverse movement during the response has a cost for the animal that depends on the total distance to choice boundary. Figure 5.8b shows that reversal probability decreases during the first 2,500 interleaved trials of the easiest decision difficulty. The combined reversal probability for the three mice was significantly higher in the first 500 trials than the final 500 trials of the period observed (15.7% vs 9.3%, Chi-square test, $p < 1e-7$).

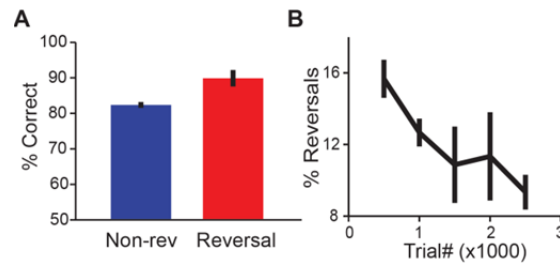


Figure 6.8: Choice reversals are used to correct errors. A: For the easiest decision category, accuracy is greater on reversal trials than on non-reversal trials averaged across all sessions for three mice. Error bars show 95% confidence intervals. B: Reversal probability for the easiest decision category decreases with task experience (average of three mice shown in blocks of 500 trials). Error bars show S.E.M.

6.5 Conclusions

We designed a simple response device for acquiring two-choice decisions from head-fixed mice. Our device has several advantages over previously used methods. By divorcing the animal's method of acquiring reward (licking) from its method of choice response (paw motion), our paradigm does not require conditioning the mouse to suppress anticipatory licking as was true for go/nogo tasks utilizing the lick response (Andermann, Kerlin et al. 2010; Schwarz, Hentschke et al. 2010). This leaves lick rate available as an additional behavioral metric. By providing high resolution trackball movement data, our device opens an additional facet of the animal's choice response to analysis that

is lost in experiments with a binary readout – the ability to detect and study changes of mind. Furthermore, the ability to reward simple lateral paw-strokes affords us greater temporal resolution than head-fixed analogs of traditional rodent decision making tasks based on directed navigation (Harvey, Coen et al. 2012).

To demonstrate the usefulness of this response device we designed an auditory psychophysics task based on binaural random click discrimination. The task was inspired by the random dot task, a simple and extensively researched paradigm in primate visual neuroscience, used to study decision making (Newsome, Britten et al. 1989; Roitman and Shadlen 2002). While we chose to use audition for superior stimulus control in head-fixed mouse, our task is fundamentally similar in that a subject is trained to integrate stochastic information supporting two hypotheses over a sampling period, and makes a discrete choice between them on each trial, providing both a choice and a reaction time measure. Our post-training sessions averaged several hundred completed trials. High trial counts are necessary to precisely characterize complex response profiles of neurons to different behavioral contingencies.

In keeping with findings in primates (Mazurek, Roitman et al. 2003; Palmer, Huk et al. 2005), mice spent more time sampling difficult stimuli (Figure 5c), and discriminated more accurately on trials with longer sampling durations (Figure 5e). Under idealized conditions in perceptual discrimination, a performance increase is realized through continued sampling of a stimulus (Link and Heath 1975; Ratcliff 1988; Mazurek, Roitman et al. 2003). Although these effects are consistent with a speed-accuracy tradeoff strategy exploiting integration of temporal evidence, further experiments specifically designed to test this hypothesis will be required (Kiani, Hanks et al. 2008; Jun, Brunton et al. 2010).

We have also shown that mice sometimes reverse the direction of choice responses they have initiated. Reversal trials correlated with performance gain, and reversal probability decreased with training. Changes of mind in human test subjects have been attributed to conflicting information in early sensory processing that is considered after a decision has been initiated (Resulaj, Kiani et al.

2009). Since after initiating a response, mice in our task were permitted to sample the stimulus continuously until their final choice was registered, the precise source of evidence used to trigger decision reversals remains unresolved. Nevertheless, our results suggest that in addition to humans, mice can also use choice reversals to strategically optimize decision accuracy. Whether the reversals we observed are true changes of mind or corrections of motor errors in support of the animal's original choice remains an open question.

The choice ball provides an opportunity to combine information-rich 2AFC psychophysical behavior with a stable, head-fixed configuration ideal for emerging imaging and optogenetic techniques. Studies in the monkeys have leveraged simple, information-rich motor responses (saccades) to uncover clearly defined set of motor control circuits in the frontal eye fields and superior colliculus (Schall and Thompson 1999), governing choice responses. A similar anatomical link has been exploited in a Go/No-go task for head-fixed mice - the ALM and PPM regions of primary motor cortex that govern licking (Komiyama, Sato et al. 2010). Since the motor response in our task is a simple lateral shift of the forelimbs, we also anticipate that these stereotyped movements will provide a tractable paradigm for understanding the neural circuits that control choice responses in mice.

Chapter 7

Conclusions and perspectives

We provided a series of findings that collectively frame confidence as a tractable problem in human and animal brain research. Beginning with a normative definition of confidence, we derived a profile of three patterns that characterize decision confidence signals. We found that these patterns generalize to remarkably different confidence computations beyond the normative model, from statistical confidence to verbal confidence reports in humans and confidence-guided investment decisions across species. To measure confidence in these studies, we designed a sensory discrimination and confidence reporting task with high resolution temporally structured evidence. The patterns in evidence preceding each confidence report belied the algorithm our subjects used to compute confidence, allowing us to resolve a family of contradicting models with a new model that uniquely reproduces a profile of robust patterns in our data. Lastly, we developed and tested a new technique for high resolution multiple choice behavior in head fixed mice, ideal for investigation of how abstract mental variables manifest mechanistically in brain circuits with optogenetic and functional microscopy techniques. In this chapter, I review each of these results in the context of their respective literatures, and conclude with implications of our confidence framework for future decision making research.

Human decision makers feel the brain's approximation of statistical confidence.

Previous research characterizing the properties of human verbal confidence has been chiefly concerned with a superficially enigmatic property of the confidence reports produced by the human organism – that these values forecast the likelihood of correct decisions astonishingly well, though slightly sub-optimally in contended ways (Chapter 1.1). However, the perspective that confidence is computed to solve a statistical problem in the deciding human brain was suggested only sparsely as a

conjecture (Kahneman and Tversky 1972; Ferrell and McGoeys 1980; Griffin and Tversky 1992), and its implications were not fully explored.

Beginning with the normative model of decision confidence (Kahneman and Tversky 1972) – that confidence is equivalent to a Bayesian posterior estimate of the probability that a decision was correct given the evidence used to decide, we formally derived three predicted interrelationships between confidence, evidence and outcome (Figure 3.1, appendix I). We showed that these relationships describe other uncorrupted signals of confidence – statistical p-values – and that this result generalized across different tests and noise distributions (Figure 3.2). The invariance of the derived patterns between Bayesian and Frequentist interpretations of confidence supported our hypothesis that these are fundamental patterns that arise when statistical confidence is computed from the evidence used to decide among options.

An intriguing suggestion we considered, is that the brain performs an operation similar to a statistical hypothesis test on the evidence it uses to make assertions, in order to derive its self-reportable feeling of confidence. We investigated the computation underlying the brain's confidence feelings by designing a two choice perceptual discrimination task, achieving very precise control over evidence delivery and behavioral measurements with Poisson click stream evidence (Sanders and Kepecs 2012; Brunton, Botvinick et al. 2013) delivered in partial acoustic isolation, by a custom, real-time pulse generator and button press capture device (section 2.1.5). In four human subjects and in an additional subject who processed our evidence in an unexpected way (Appendix III), the interrelationships between three measures of each decision –choice outcome, the strength of evidence and confidence, robustly exhibited the same characteristic patterns which arise in both Bayesian and Frequentist statistical confidence measures.

Since confidence is an operation performed on samples of evidence, we sought to determine whether this result was dependent on something special about samples of evidence provided by the brain's auditory perceptual pipeline. Specifically, we wondered whether normative patterns would persist in confidence, if the samples of evidence used to decide were gathered from memory (Ratcliff

1978). We tested 21 subjects providing 100 or 150 trials each, in a two choice task where subjects compared quantities in long-term memory that were general knowledge – the populations of countries - and reported their confidence in their decisions. Normative patterns robustly described confidence in this task as well, and emerged when individual subjects provided as few as 100 trials (Figure 3.14), confirming that extensive training on the statistics of a task was not necessary to observe these patterns.

We conclude that the human *feeling* of decision confidence exhibits properties that are true of all *statistical* decision confidence measures. However, we cannot exclude the possibility that this profile of patterns could also arise from a signal that is in some way differentiable from decision confidence – a topic I explore in appendix IV by contrasting the patterns produced by an evidence gain modulation factor against the patterns of statistical confidence in the three normative model projections. In the context of this comparison, I suggested that a signal containing the three normative patterns is most likely as useful to a decision maker as a confidence measure for predicting outcomes and casting wagers – and that any further difference between the signal and statistical confidence is likely a semantic problem. While the wide range of confidence measures assayed in this dissertation establishes that the *sensitivity* of these patterns for things we call confidence is high, further research is necessary to determine the *specificity* of these patterns as a putative identity test for statistical confidence signals.

Our results were consistent with several well established properties of human decision confidence – that decision confidence strongly forecasts outcomes in sensory discrimination tasks (Keren 1988; Björkman, Juslin et al. 1993; Baranski and Petrusic 1994; Olsson and Winman 1996; Merkle and Van Zandt 2006), that confidence is correlated with the strength of provided evidence (Jastrow and Peirce 1884; Vickers and Packer 1982; Baranski and Petrusic 1998) and that confidence in perceptual tasks without speed instructions, is anti-correlated with reaction time (Volkman 1934; Merkle and Van Zandt 2006). Our results from Chapter 3 also establish a *new* pattern in human verbal confidence reports – that the lowest verbal confidence occurs in the presence of the errors informed

by the strongest evidence. While we observed this pattern in two markedly different behavioral tasks, further research will be required to establish whether this is truly a general pattern that arises from the nature of imperfect evidence sampling, or whether the directions of the confidence/evidence relationships can be changed in humans and animals by different task contingencies.

Finally, and speculatively, the fact that reports from humans were verbal carries the curious implication that reported feelings may provide conscious access to brain's approximation of statistical computations. Common functional and circuit architectures in the brain's mechanistic underpinnings of those unitary mental variables that can be "felt" and directly reported, may eventually belie the purpose of conscious experience in organisms.

Confidence informs time investment in two distant mammalian species.

In 2008, Kepecs et al. discovered that when considered with respect to evidence and choice outcome, the firing of individual neurons in the rat orbitofrontal cortex produce the same patterns we have attributed to statistical confidence in chapter 3 (Kepecs, Uchida et al. 2008). To further study the neural correlates of confidence in rodents, we needed a non-verbal way to knowingly attribute the human and statistical notions of confidence to behavioral measures in animals. Previous methods had relied on the probability that an animal chooses an "uncertain" option for a small reward in a two-choice task (Watson, Kellogg et al. 1973; Smith, Shields et al. 1997), or declines the opportunity to answer a given difficult choice for a small reward (Foote and Crystal 2007; Kiani and Shadlen 2009). Neither of these behaviors have been directly compared with explicit confidence in humans. More importantly, these confidence measures are binary, prohibiting the kind of detailed analyses we presented in chapter 4, and do not provide within-trial confidence measures that are necessary to observe the relationship between the discriminability of a choice and subsequent confidence.

To address this problem, we developed a new task where subjects sampled sensory evidence (Poisson clicks) and *invested time* in the hypothesis that they were correct. We found that time investments generated the trend directions of all three normative model patterns expected of statistical

confidence in both humans (Figure 4.3) and rats (Figure 4.7), though deviated from normative and verbal confidence by a non-linear transform. In humans, we requested an additional confidence measure after each time investment: an explicit confidence report. This provided the first within-trial comparison of implicit and explicit confidence measures, to the limit of our awareness. We reported a strong correlation between verbal confidence reports and time investment, and countered the possibility that time investment was the source of these verbal reports (Figure 4.5).

We thus provide two lines of clear evidence supporting our hypothesis that time investments were derived from decision confidence. 1. The presence of robust normative model patterns in time investment (Figure 4.3b, 4.7c), implies that the process which computed time investment had access to the perceptual evidence used to make decisions. 2. The strong correlation we observed between time investment and verbal confidence is consistent with the hypothesis that these implicit and explicit reports were derived in series from the same internal confidence value on each trial.

By establishing that time investment shows the hallmarks of a confidence-derived measure and co-varies with verbally reported confidence in humans, we have provided an empirical basis for interpreting implicit confidence in humans and rodents. Subsequent research may seek to identify the neural circuits underlying confidence computations in the brain using the rodent model.

It is unknown whether the alternative processes of mentally observing confidence to report it verbally, or instead, using confidence implicitly to guide behavior, affect the way confidence manifests. Future studies which are designed to measure a subject's cost functions for time and reward utility using our task, may seek to recover the subject's abstract confidence prior to its use in computing time investment (Equation 4.1), to determine whether its form is affected by its intended function, and how it varies across conditions. A phenomenological understanding of the basic properties that govern how the brain computes and uses this important mental variable will clarify its potential roles in different behaviors and disorders that rely on learning and mitigation of risk.

The human brain's algorithm for sensory discrimination and post-decision confidence resembles a coupled accumulator adapted to exploit noisy evidence gathered after choice commitment.

Sequential sampling models of choice and confidence have sporadically appeared over the past half-century, positing conflicting algorithms for deriving confidence from evidence. To determine which model could account for human confidence, we designed a confidence task with unprecedented resolution in the timing of patterned evidence delivery and response acquisition, to capture signatures of the algorithm subjects used to decide.

Among five previously posited computational models, the OCA model we developed is unique in its ability to simultaneously account for both patterns in evidence history and the signature patterns of statistical confidence. It further provides a “global” fit that reproduces most qualitative aspects of the interrelationships between choice, reaction time, discriminability, confidence and evidence history in the 12 key projections simultaneously assayed. While it does not perfectly model reaction time under deadline stress, this simple model constitutes a significant step towards a general algorithmic understanding of how the brain computes choice and confidence.

Two-choice psychometric discrimination in head-fixed mice can be readily achieved by conditioning lateral paw motion, captured by our Choice Ball device.

Precise two alternative forced choice (2AFC) behavior was not previously available in head-fixed mouse, complicating investigation of the brain's decision making variables with technologies that require imaging to capture and manipulate aspects of brain function. To realize 2AFC behavior in this constraining setting, we designed a simple response device enabling mice to report two-choice decisions with paw movements, and trained mice to discriminate Poisson click stream pairs of varying difficulties. We showed that head-fixed mice can learn the same psychometric discrimination behavior that we used for our human subjects and rats to study confidence, and respond on average with rapid (~180ms) paw responses, producing psychometric and chronometric signatures of a well-

controlled psychophysical behavior. This finding may provide a path for future research targeting the neural circuits underlying decision confidence, with appropriate modifications to capture an effort investment or time investment measure in addition to choice. We anticipate that paw response measures in head-fixed mice will be exploited in future studies of decision making to simultaneously leverage the advantages of an informative decision making behavior and advanced genetic and optical imaging tools.

Implications of our framework

The chapters of this thesis form a series of conceptual bridges between computational, algorithmic and mechanistic levels of understanding confidence, which can be leveraged as a framework to guide design of future investigations. We began with a first-principles definition of confidence - in essence, a statistical computation of likelihood. We derived three signature patterns of this computation, and showed that both statistical p-values and the human feeling of confidence bear striking resemblance to these patterns, while other measures lacking information about the decision maker's internal evidence, do not. We take from this, though with due caution, the central insight of our framework - that the fundamental operation underlying the brain's sense of confidence is a computation approximating statistical likelihood.

Alongside inferences from behavior, further insight will come from observing the computation of confidence directly, as it manifests in the brain. Towards this end, we reported signatures of a confidence calculation in human and rodent time investments, and supplemented our argument that these are products of the brain's use of confidence by showing a direct correlation between time investment and verbal confidence. This finding includes within our framework, a convenient definition of *applied* confidence in behavior, and methods by which it can be reproduced in model organisms ideal for brain research.

We leveraged insight from the computational level of our framework to resolve the *algorithm* of confidence, by requiring that it account for all three normative model patterns in human data. The OCA model further adds to our framework, providing a realistic way to realize the computation of confidence in real-time. This algorithmic understanding forms a second bridge, permitting the computation of confidence to be understood in terms of neural circuit mechanisms, which also process information in real-time and cannot accept abstract symbols as inputs.

To study the brain's internal representation of confidence in terms of neural circuits, our framework provides two useful tools – an algorithm and a class of behavior. We added to this, a partial implementation of this behavior in a manner compatible with functional microscopy and ideal for optogenetic dissection of circuit function.

We anticipate that the combined tools and insights of our framework will endow researchers with unprecedented access to the physical correlates of the familiar statistical experience we know as confidence.

Appendix I

Derivation of confidence signal properties from normative confidence

The following three derivations were courteously contributed by Dr. Balazs Hangya, presently a postdoctoral fellow in our research group, to determine what properties are expected of confidence signals in their interrelationship with choice accuracy and discriminability.

Confidence is correlated with choice accuracy

Arbitrating between opposing hypotheses and assigning levels of confidence to our decisions can be viewed as a statistical hypothesis testing problem. We make a decision based on an internal variable (decision variable, \hat{D}), which is the internal reflection - estimate - of a corresponding external variable (D). (Remark: D and \hat{D} can be multidimensional.)

Definition 1: Let us denote the external variable D and realizations of this random variable d . Let us denote the corresponding internal variable \hat{D} and realizations of this random variable \hat{d} (referred to as *percept*; also called the *decision variable*). We define another random variable called the *choice* (also termed *decision*), denoted by θ (realizations will be denoted by ϑ). The choice is a probabilistic function of the percept. Thus, a realization of the choice takes the form of a probabilistic event of $\hat{d} \in \theta$, where θ is subset on the space of all percepts (*percept space*). (Because this represents an equivalent description of the choice, we keep the same Greek letter for the notation; however, we use bold face rendering whenever θ refers to a subset.)

Our decision can be thought of as the choice of a hypothesis (in statistical terms, the alternative hypothesis) against all possible complementary choices (constituting the null hypothesis):

null-hypothesis (H_0): $\hat{d} \in \theta_0$

alternative hypothesis (H_1): $\hat{d} \in \theta_1$

where θ_1 represents our choice. The choice is called *correct* if the alternative hypothesis is true and *incorrect* otherwise. In this context, our confidence (c) can be defined as the probability of the alternative hypothesis being true provided the percept and the choice.

Definition 2: Let *confidence* be defined by

$$c = P(H_1 | \hat{d}, \vartheta)$$

(As usual, the random variable will be designated by C and its realizations by c). We can thus define the function determining confidence from the decision variable and the choice: let the function $\xi: \mathcal{R}(\theta) \times \mathcal{R}(\hat{D}) \rightarrow [0,1]$ be defined by

$$\xi(\vartheta, \hat{d}) = P(H_1 | \hat{d}, \vartheta)$$

where $\mathcal{R}(\hat{D})$ denotes percept space and $\mathcal{R}(\theta)$ denotes the range of all possible choices (i.e., the *choice space*).

Next, we will derive the relationship between confidence and accuracy.

Definition 3: *Accuracy* is the expected proportion of correct choices.

We seek to determine the following function:

$$f: [0,1] \rightarrow [0,1]$$

$$f: c \mapsto A_c$$

where A_c is the accuracy for choices with a given confidence. Our claim is that this function is identity.

Theorem 1: Accuracy equals confidence:

$$A_c = c$$

Proof: For every given level of confidence, there is a set of values of percept-choice pairs leading to the same confidence value: let us denote the image of c by ξ^{-1} as $\{(\hat{d}_i, \vartheta_i)\}_{i \in I}$, the set of choice-decision variable pairs mapping onto c . Let us first assume that I is a countable set. Accuracy for confidence level c is determined by the probability of a correct choice if $C = c$ over the probability of encountering the confidence level of c (that is, $P(C=c)$):

$$A_c = \frac{\sum_{i \in I} P(H_1, \hat{d}_i, \vartheta_i)}{\sum_{i \in I} P(\hat{d}_i, \vartheta_i)}$$

From the definition of joint probability,

$$A_c = \frac{\sum_{i \in I} P(H_1, \hat{d}_i, \vartheta_i)}{\sum_{i \in I} P(\hat{d}_i, \vartheta_i)} = \frac{\sum_{i \in I} P(H_1 | \hat{d}_i, \vartheta_i) \cdot P(\hat{d}_i, \vartheta_i)}{\sum_{i \in I} P(\hat{d}_i, \vartheta_i)}$$

As we know that $\forall i \in I: P(H_1 | \hat{d}_i, \vartheta_i) = c$,

$$A_c = \frac{\sum_{i \in I} P(H_1 | \hat{d}_i, \vartheta_i) \cdot P(\hat{d}_i, \vartheta_i)}{\sum_{i \in I} P(\hat{d}_i, \vartheta_i)} = \frac{\sum_{i \in I} c \cdot P(\hat{d}_i, \vartheta_i)}{\sum_{i \in I} P(\hat{d}_i, \vartheta_i)} = \frac{c \cdot \sum_{i \in I} P(\hat{d}_i, \vartheta_i)}{\sum_{i \in I} P(\hat{d}_i, \vartheta_i)} = c$$

I is not necessary a countable set. We can re-write the equations in continuous form to apply to any sets as follows.

$$A_c = \frac{\int_{\Pi=1} \int \int_{(\hat{D}, \theta) \in \xi^{-1}(c)} f_{\Pi, \hat{D}, \theta}(\pi, \hat{d}, \vartheta) d\vartheta d\hat{d} d\pi}{\int \int_{(\hat{D}, \theta) \in \xi^{-1}(c)} f_{\hat{D}, \theta}(\hat{d}, \vartheta) d\vartheta d\hat{d}}$$

where Π is a random variable that is 1 if the choice is correct, and 0 otherwise.

$$\begin{aligned}
A_c &= \frac{\int_{\Pi=1} \int_{(\hat{D}, \theta) \in \xi^{-1}(c)} f_{\Pi, \hat{D}, \theta}(\pi, \hat{d}, \vartheta) d\vartheta d\hat{d} d\pi}{\int \int_{(\hat{D}, \theta) \in \xi^{-1}(c)} f_{\hat{D}, \theta}(\hat{d}, \vartheta) d\vartheta d\hat{d}} \\
&= \frac{\int_{\Pi=1} \int_{(\hat{D}, \theta) \in \xi^{-1}(c)} f_{\Pi}(\pi | \hat{D} = \hat{d}, \theta = \vartheta) \cdot f_{\hat{D}, \theta}(\hat{d}, \vartheta) d\pi d\vartheta d\hat{d}}{\int \int_{(\hat{D}, \theta) \in \xi^{-1}(c)} f_{\hat{D}, \theta}(\hat{d}, \vartheta) d\vartheta d\hat{d}} \\
&= \frac{\int_{\Pi=1} \int_{(\hat{D}, \theta) \in \xi^{-1}(c)} c \cdot f_{\hat{D}, \theta}(\hat{d}, \vartheta) d\pi d\vartheta d\hat{d}}{\int \int_{(\hat{D}, \theta) \in \xi^{-1}(c)} f_{\hat{D}, \theta}(\hat{d}, \vartheta) d\vartheta d\hat{d}} = \frac{c \cdot \int \int_{(\hat{D}, \theta) \in \xi^{-1}(c)} f_{\hat{D}, \theta}(\hat{d}, \vartheta) d\vartheta d\hat{d}}{\int \int_{(\hat{D}, \theta) \in \xi^{-1}(c)} f_{\hat{D}, \theta}(\hat{d}, \vartheta) d\vartheta d\hat{d}} = c
\end{aligned}$$

□

Remark 1: We would like to note that these consideration about confidence do not depend on a particular theory of perception, that is, the function mapping the external variable on the internal percept: $D \mapsto \hat{D}$. Furthermore, the derivations also do not depend on a particular theory of decision, that is, the function between the internal variable or percept and the decision or choice: $\hat{D} \mapsto \theta$. This includes both deterministic and stochastic decision models, the latter referring to models where a certain percept can result in more than one decision. In case of deterministic decision models, the percept unequivocally determines the choice, thus in the equations we could drop the choice from the inverse picture of a confidence value, taking only the percept into account: $\{\hat{d}_i\}_{i \in I}$ instead of $\{(\hat{d}_i, \vartheta_i)\}_{i \in I}$. However, as this simplified version would not include stochastic decision models, we chose to adhere to the general formalization.

Remark 2: We also note that there is no need for a relation to be defined on the percept space. However, if the choice is fixed (or determined by the percept, as in deterministic decision models), confidence defines a natural relation on percepts by ξ . More precisely, the order relation on confidence values can be pulled back to the percept space by taking $\xi^{-1}(c)$ and restricting it to a particular choice.

Confidence is correlated with discriminability (difficulty) on correct trials, and anti-correlated on incorrect trials.

Perceptual decisions involve a feeling of decision difficulty. Next, we will formally introduce the concept of difficulty and examine the changes of confidence along this difficulty axis.

Definition 4: Let us define choice *difficulty* as the probability of evidence for incorrect choice:

$$P(\hat{d} \in \theta_0)$$

Remark 3: Please note that this definition is although conceptually useful, it will not be strictly applied in the following derivations and hence other definitions are also possible without

impacting the overall results. For deterministic decision models, the proportion of errors will equal the above probability. In the following theorem, the monotony assumption (second assumption, see below) is in agreement with the above definition; however, they are not consequences of one another.

Theorem 2: Let us assume that

- ξ is independent of difficulty
- the distribution of percepts changes monotonically with difficulty

Under these assumptions, confidence increases for correct choices and decreases for incorrect choices with decreasing difficulty.

Proof: We start with the somewhat counterintuitive claim regarding the incorrect choices. Let us first examine the two assumptions in more detail.

The first assumption postulates that the function from percept to confidence does not change with difficulty. Thus, whenever we calculate expected value of confidence over a percept distribution, only the percept distributions will depend on difficulty.

For incorrect choices, the second assumption means that with difficulty decreasing, the relative weight of low-confidence percepts increases while the relative weight of high-confidence percepts decreases in the percept distribution. Note that low-confidence and high-confidence percepts are defined here through the relation imposed by ξ on the percepts; see *Remark 2* in the previous section. As a trivial consequence of this definition, confidence changes monotonically along low- and high-confidence percepts.

Let us consider two different levels of difficulty (S_1 and S_2), with corresponding distributions of percept restricted to incorrect choices P (S_1 , difficult) and Q (S_2 , easy). It is sufficient to show that the expected value of confidence is larger for S_1 than for S_2 :

$$\int_0^1 c \cdot p(c)dc > \int_0^1 c \cdot q(c)dc$$

where p and q denotes the probability density functions corresponding to P and Q , respectively. Note that $p(c)$ can be thought of as the probability of the picture of c by ξ^{-1} restricted to incorrect choices in the percept space.

Equivalently,

$$\int_0^1 c \cdot [p(c) - q(c)]dc > 0$$

Let us denote $I_0 \subset [0,1]$ the interval where $p < q$, and $I_1 \subset [0,1]$ the complementary interval where $p > q$. The existence of these intervals is the consequence of the monotony assumption. Thus, there is a critical confidence value (denoted here by c_{crit}) for which $I_0 = [0, c_{crit}]$ and $I_1 = [c_{crit}, 1]$. We then re-write confidence as $c = c_{crit} - c'$ if $c < c_{crit}$ and $c = c_{crit} + c'$ if $c > c_{crit}$; thus, $c' > 0$ for both cases. Applying these notations,

$$\begin{aligned}
& \int_0^1 c \cdot [p(c) - q(c)] dc = \\
& = \int_0^{c_{crit}} c \cdot [p(c) - q(c)] dc + \int_{c_{crit}}^1 c \cdot [p(c) - q(c)] dc = \\
& = \int_0^{c_{crit}} (c_{crit} - c') \cdot [p(c_{crit} - c') - q(c_{crit} - c')] dc' + \\
& \quad + \int_0^{1-c_{crit}} (c_{crit} + c') \cdot [p(c_{crit} + c') - q(c_{crit} + c')] dc' = \\
& = c_{crit} \cdot \left(\int_0^{c_{crit}} [p(c_{crit} - c') - q(c_{crit} - c')] dc' + \int_0^{1-c_{crit}} [p(c_{crit} + c') - q(c_{crit} + c')] dc' \right) \\
& \quad + \int_0^{1-c_{crit}} c' \cdot [p(c_{crit} + c') - q(c_{crit} + c')] dc' \\
& \quad - \int_0^{c_{crit}} c' \cdot [p(c_{crit} - c') - q(c_{crit} - c')] dc' = \\
& = c_{crit} \cdot \left(\int_0^{c_{crit}} [p(c) - q(c)] dc + \int_{c_{crit}}^1 [p(c) - q(c)] dc \right) \\
& \quad + \int_0^{1-c_{crit}} c' \cdot [p(c_{crit} + c') - q(c_{crit} + c')] dc' \\
& \quad + \int_0^{c_{crit}} c' \cdot [q(c_{crit} - c') - p(c_{crit} - c')] dc' > 0
\end{aligned}$$

In the last step, the first term is 0, since

$$\begin{aligned}
& \int_0^{c_{crit}} [p(c) - q(c)] dc + \int_{c_{crit}}^1 [p(c) - q(c)] dc = \\
& = \int_0^1 [p(c) - q(c)] dc = \int_0^1 p(c) dc - \int_0^1 q(c) dc = 1 - 1 = 0
\end{aligned}$$

The second term is positive, since c' is positive and the probability density functions are evaluated on I_1 , where $p > q$. Finally, the third term is also positive, because c' is positive and the probability density functions are evaluated on I_0 , where $q > p$. In consequence, the sum is positive, which concludes the proof for incorrect choices.

For correct choices, the high-confidence percepts are increasingly more likely with decreasing difficulty, thus showing an opposite pattern as compared to incorrect choices. Therefore, a symmetric derivation proves the increase of confidence with decreasing difficulty for correct choices.

Remark 4: The first assumption is essential. If confidence is defined based on the relevant conditional probability (see *Definition 2* in the first section) from a distribution reflecting a single difficulty level, then the function between the percept and confidence (ξ) will differ among difficulties, thus the above proof does not apply. Furthermore, the expected value of confidence cannot decrease with decreasing difficulty for incorrect choices, for the following reason. For maximal difficulty, when the choice is random with respect to the percept, confidence will fall to its lowest possible value (0.5), reflecting equal probabilities of the null and alternative hypothesis regardless of the percept. This translates to situations in which a person (or other decision making agent) is provided with information about difficulty, e.g. by grouping decisions of similar difficulties, giving a chance to learn about difficulty and define confidence accordingly. Thus, the above theorem only applies when changing confidence levels based on knowledge of choice difficulty is prevented, e.g. by randomizing the order of choices with different difficulty.

For a fixed discriminability, confidence predicts choice outcome.

Corollary 1: For any given difficulty, accuracy for low confidence choices is not larger than that of high confidence choices (dividing the confidence distribution at any particular value). Strict inequality holds in all cases when accuracy is dependent on the percept.

Proof: Let us take the set of low-confidence percept-choice pairs corresponding to the low confidence choices by ξ^{-1} , and similarly, the set of high-confidence percept-choice pairs corresponding to the high confidence choices. By the definition of confidence (*Definition 2* in the first section), low-confidence percept-choice pairs cannot have higher accuracy than the high-confidence percept-choice pairs. If all percepts are associated with the same accuracy (either when the percept does not carry information about the hypotheses of choice, or when the percept determines the correct choice with a probability of one), the two accuracies are equal. Otherwise, the two accuracies should necessarily differ, in which case the strict inequality holds.

Appendix II

Bpod: a microcontroller based real-time behavior acquisition and control system

Modern research on the biological basis of cognition builds heavily upon the successes of the behaviorist tradition, underscoring the power of precise and reliable behavioral metrics as a proxy for mental processes. The earliest of modern animal behavior studies used experiment-specific analog devices to achieve behavioral precision through automation (Skinner 1938). Automation of behavior was simplified by the later use of specialized personal computer software to design and execute behavior protocols (Brainard 1997; Peirce 2007), though largely at the expense of the ability to control behavior with millisecond precision (Ramamritham, Shen et al. 1998). The need for this level of precision is inspired by the insight that sub-millisecond action potentials in cell populations are a fundamental currency of mental information (Hebb 1949). Ideally, in modern experiments relating the firing of neurons to behavior, behavioral measures and manipulations should match the brain's fundamental signals in their precision. Historically, this ideal has been compromised in favor of convenient instrumentation.

Use of an operating system with a real-time kernel and low-level programming of behavioral protocols can provide sub-millisecond precision using a PC, but the advanced programming skills required restrict this approach. B-control (C.D. Brody, Z.F. Mainen, A.M. Zador, CSHL) is a rodent behavior system designed to address this problem, and has been used in several published studies (Felsen and Mainen 2008; Erlich, Bialek et al. 2011; Brunton, Botvinick et al. 2013; Znamenskiy and Zador 2013). B-control allows trials in behavioral tasks to be organized in MATLAB (Mathworks) as a matrix of states and state transitions, and delegated to a separate real-time computer for execution. Under the governance of a trial's state matrix, B-control provides the ability to precisely read photogates, drive solenoid valves, deliver acoustic stimuli and respond to logic signals.

In this section, I advance a behavior system that leverages a microcontroller platform to replicate most of B-control's functionality, and builds upon B-control's central idea in several ways

that improve generality and accessibility. In section 4.3.5, we use data acquired with this new system to show several classic psychophysics measures and confidence usage in rats performing the Poisson clicks task.

Rationale for an inexpensive open hardware state machine

While B-control has proven to be an excellent system for capturing rodent behavior, we identified three major aspects that could be improved. 1. Modifying the system to interface with new actuators and sensors requires programming in C for real-time Linux. Ideally, the system would be implemented on a platform more accessible to researchers without education in low level programming. 2. The B-control system relies on proprietary, custom-designed hardware, only available from a specific freelance engineer. Access to the system depends upon the availability of that engineer to provide the hardware. 3. Since the system uses a dedicated computer to run the state machine, a proprietary PC interface card and a custom interface module, a single installation costs thousands of dollars.

One way to achieve real-time precision in behavior in a way that is more technically and financially accessible is to take advantage of the growing popularity of open source microcontroller platforms. Arduino is one such platform which has previously been applied to a range of applications in Neuroscience research (Sun, Bouchard et al. 2010; Teikari, Najjar et al. 2012) and leverages thorough documentation with simple tutorials alongside a large community of active developers (available at <http://arduino.cc/>). While Arduino provides the precision necessary for a real-time behavior solution, it lacks peripheral interfaces to directly control instruments and read sensors used in animal behavior. Also, native Arduino restricts behavior protocol development to a modified subset of c++, complicating protocol writing for programmers inexperienced with low-level performance optimization for real-time applications.

To address these limitations, I developed Bpod - a device that uses an Arduino microcontroller to implement the abstraction of real-time precision for high level state machine-based

behavioral protocols. To interface between the microcontroller and experiment, a custom designed printed circuit board provides a convenient hardware bridge to common equipment for manipulating a rodent's environment – LEDs, TTL logic lines and solenoid valves. The same circuit board provides an interface for behavior measurement using infrared photo-gates, analog sensors and TTL trigger lines. To further expand Bpod's usefulness, we have expanded Bpod to interface with both Ethernet devices and an open source programmable pulse generator for laser control in optogenetics and control of patterned stimuli like our auditory Poisson click stimulus (Pulse Pal, section 2.1.6). Firmware on the microcontroller executes a virtual state machine with guaranteed 300 μ s resolution on generated and captured events, and seamlessly interfaces with the MATLAB client using a virtual serial port. A single installation can be assembled for under \$250 using common parts from industrial suppliers, raising the possibility of high throughput animal training in environments with limited funding. By combining an augmented open source microcontroller with powerful developer tools for achieving sub-millisecond control of behavioral experiments, we anticipate that our system will be useful for a variety of applications ranging from in vivo electrophysiology to psychophysics, behavioral phenotyping and science education.

Circuit and firmware design

The schematics, bill of materials, enclosure CAD files and MATLAB client software are open source, and available at the following URL: <https://bitbucket.org/kepecslab/bpod>. This section is intended to provide a very brief, non-technical explanation of how the circuit and firmware operate.

Bpod is built on the Arduino Mega 2560 platform, a programmable single-chip computer (microcontroller) with added power and USB communication peripherals. To interface between Arduino Mega and the operant chamber's sensors and actuators, I designed the circuit board drawn in Figure A2.1. This figure is not intended to show connectivity, but to pictorially introduce the major components of the circuit.

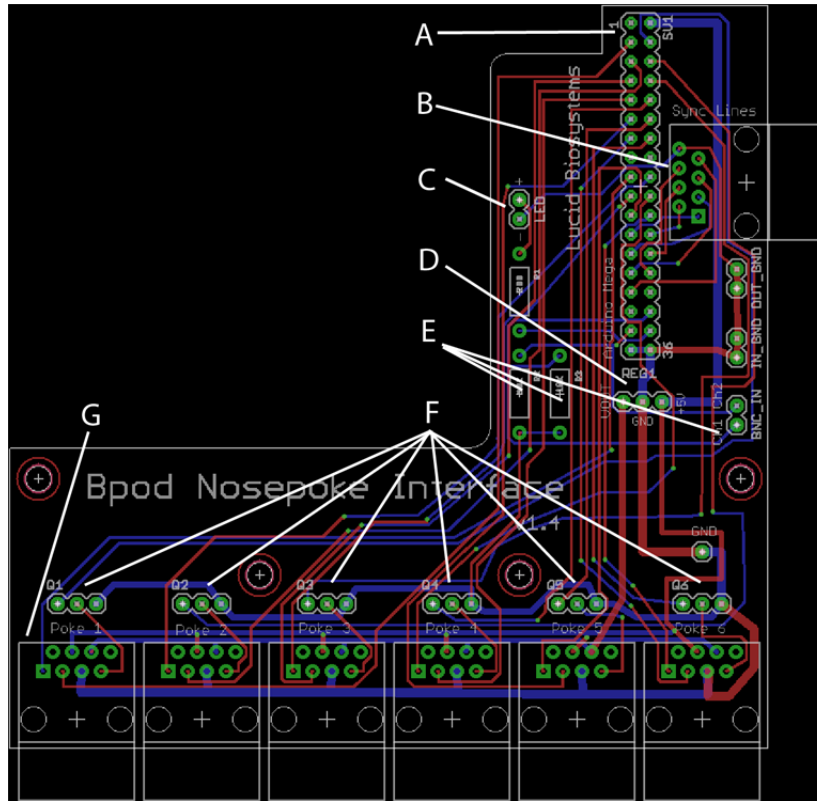


Figure A2.1: Components of the Bpod interface circuit board. **A:** 36 pin plug to connect Arduino. The top two pins as shown provide the USB 5V supply, and bottom two provide USB ground. All other pins are bidirectional 5V digital logic lines. **B:** 8 pin RJ45 Ethernet jack to synchronize behavior states with an electrophysiology system. One pin is a ground pin. Using the 7 remaining pins, up to 128 states can be synchronized using a parallel binary code. **C:** Indicator LED remains on while Bpod is running a state matrix. **D:** A 2 Amp boost regulator (Pololu), converts the USB 5V supply to 12V for powering solenoid valves. With this regulator, a laptop can power the entire behavior system, making it robust against power fluctuations. **E:** Input BNC signal connection pins and pull-down resistors. Pull-down resistors force the signal lines to ground potential until a high current (TTL) signal arrives. **F:** logic-gated MOSFET transistors allow Arduino's native logic signals to connect and disconnect solenoids and their 12V power source from ground for precision liquid and gas delivery. **G:** RJ45 plug for nose port interface (1 of 6 indicated). Red lines indicate electrical connections on the top face of the board, and blue lines indicate bottom-side connections. Green circles indicate connections between the top and bottom layers.

The firmware is software written in c++, which is converted by Arduino's software tool-chain to AVR assembly language, and written to the microcontroller's program flash memory. The microcontroller does not have an operating system. It stores and runs only the program intended. This way, the Bpod firmware has total control of the time course of events. The firmware is a set of instructions that are run indefinitely in a loop. Pseudo-code for the loop is shown in figure A2.2.

```

Loop start:
  Wait for an op code to arrive from the USB connection.
  If the op code is:
    1, This was a HANDSHAKE. Send back an acknowledgement byte.
    2, This was a request to OVERRIDE an I/O line's logic value (0V or 5V).
      Read 2 bytes - the line ID and the intended state. Change the line voltage.
    3, This is a request for Bpod to accept a new state matrix. The matrix arrives
      as a long string of bytes, which must be re-assembled into a matrix in the
      microcontroller's memory. Long data types such as 32-bit timer values
      are stored in separate arrays to conserve memory. Subsequent bytes
      indicate whether the matrix uses peripherals that need to be initialized
      (i.e. Ethernet). If necessary, initialize them.
      Send back an acknowledgement byte that the matrix was received.
    4, This is a command to RUN the state matrix that is currently stored.
      Adjust all I/O lines to match the first state in the matrix.
      Read the current states of photogates and BNC lines.
      Read the start time of the trial and current state from the microcontroller clock
      Set MatrixFinished = 0.
  While MatrixFinished == 0, run the following loop:
    Check to see if a new byte has arrived from the USB port. If so,
    If the byte is:
      5, This is a request to OVERRIDE an I/O line. Read bytes and execute.
      6, This is a VIRTUAL POKE. Set variables to emulate the intended port state.
      7, This is a software force quit. Set MatrixFinished = 1.
    end
    Read the current states of photogates and BNC lines.
    Check current states against last known states.
    Set the current event code equal to 0 (no event)
    If the state of one of these has changed,
      Store the new event code (same as column of state matrix)
    end
    Read the current time
    Check whether an active scheduled wave elapsed. If so inactivate & set event code
    Check whether the current state timer elapsed. If so, set event code.
    If the event code's column in the current state row of the matrix indicates a state change,
      Update the current state according to the state matrix
      Write the new state number in binary to the electrophysiology sync lines
      Set I/O lines to values determined by the new state (LEDs, valves, BNC, etc)
      Write new state values to peripherals (i.e. Ethernet)
      Activate any scheduled waves started by new state, update end time variables.
    end
    If the event code is not 0
      Store the timestamp for this event by appending to the timestamp array
    end
    Send the event code to the Bpod client by USB (1 byte)
    If the new state is state number 0 (the exit state),
      set MatrixFinished = 1.
    end
  end
  Send event code 255 to the Bpod client (tells the client to expect timestamps)
  Return 32-bit timestamps for all events via USB
  Terminate all active scheduled waves
end
end

```

Figure A2.2: Pseudo-code of Bpod firmware. This is an English readable description of the firmware code available at <https://bitbucket.org/kepecslab/bpod>. It does not include variable declarations or other initialization routines that are run when the Bpod device is first powered.

Software client features

To expedite behavior protocol development and execution, the Bpod client for MATLAB contains an array of software tools including a state matrix assembler for simple text description of state matrices, a data manager for organizing each subject's data and settings, and a solenoid valve

calibrator to simplify precise liquid and gas delivery. A higher level protocol development tool is available to save time on coding when creating GO/NOGO or 2AFC protocols.

2.3.4 Conclusions

The Bpod system constitutes a small step towards an ideal system for rodent behaviors commonly used in Neuroscience research. It is inexpensive, reliable and has been chosen to power projects ranging from the presently reported confidence task to conditioned place preference, self-stimulation, head-fixed auditory GO/NOGO, rat soccer and social value measurement. However, several challenges remain to be solved before Bpod is adopted more broadly. For tasks that require high fidelity auditory cues instead of simple clicks and feedback tones, Pulse Pal is an insufficient solution for sound delivery. Also, Bpod's software client lacks modular plugin architecture, for conveniently adding arbitrary sensors and actuators. This can be added in a future update. Finally, Bpod's onboard 8kb of RAM is only sufficient to store matrices of 50 states and to record 1,000 events per trial as currently programmed. Though we have not yet exceeded these limits in our lab, constraints are not ideal. The memory limit has been addressed in a new design available in the repository, based on a faster microcontroller with 9x more memory (Arduino Due). This new prototype has been bench-tested, but it has not yet been used in research.

While the system has already proven a valuable asset, we anticipate that with the improvements above, Bpod will become the interface of choice to leverage powerful microcontrollers for real time precision in behavioral research.

Appendix III

Decision confidence with an alternative evidence processing strategy: a case study

In section 3.1.2, we noted that for one of our subjects (H5), the first side to click predicted choice on 74% of evidence-neutral trials. However, the subsequent 7 clicks also predicted choice with greater than chance accuracy (Figure A3.1a), indicating that subsequent information was considered. The subject's mean reaction time was 406ms, (Figure A3.1b), far closer to the 294ms mean response time of our rats than the reaction times of our other human subjects. We wondered whether normative model patterns would persist in explicit confidence reports, in a regime where processing of evidence occurred on a short timescale. To observe confidence in this subject at high resolution, we gathered a large dataset (11,463 trials over 18 sessions).

The period of information that contributed to subject H5's average choice lasted for only 250ms after a 150ms motor response delay (Figure A3.1c), indicating that at 100 clicks per second, 24 clicks on average *could* have been considered besides the first click. Most significantly for our analysis, the reverse correlation shows that anti-chosen stream evidence was selected against in making choices almost as much as chosen stream evidence was selected for, indicating that despite over-contribution of the first click, the *balance* of evidence was the dominant feature driving choice.

Consistent with this view, a psychometric curve fit with a binomial logistic function as in Figure 3.9 indicated that our balance of experienced evidence measure, β' , still meaningfully described choice probability (Figure A3.1d), achieving 100% choice accuracy on both sides. A slight side bias (same panel) and a considerable feedback bias (Figure A3.1e) also contributed to choice. We determined that despite short response times, subject H5 sampled for longer on difficult trials, and exhibited reduced accuracy with prolonged sampling as for our other subjects (Figure A3.1f-g).

Short response times at such high accuracy were possible in part, because we gave the subject strong evidence to compensate for low initial performance. Taken together, the subject accumulated strong evidence on a short timescale, selectively prolonged sampling of ambiguous evidence, and at

least partially evaluated its balance to make choices. Since our metric of evidence meaningfully described choice, we sought to determine how well it described confidence in a regime prone to several biases.

Subject H5 responded with a confidence frequency distribution unlike any of the subjects in our previous report (Figure A3.1h), using the “5” response more often than the remaining parts of the scale combined. Despite this unusual scale usage, confidence reports reproduced the three normative model patterns, capturing the correct slope directions for all trends (Figure A3.1i-l). Calibration (section 1.1.1) was nearly perfect at 0.003, indicating that the subject was a good forecaster of her outcome probabilities. The relationship of confidence to evidence exhibited the expected correlation and anti-correlation for correct and incorrect trials, however the correct trial data saturated at 5 on very strong evidence trials, making the curve a poor linear fit. A similar saturation exists for the p-value of the one-sided bootstrap confidence test for difference between means, using Poisson evidence (Figure 3.2h). The subject’s confidence had a remarkable ability to discriminate between correct and incorrect trials of the same experimentally delivered evidence strength (Figure A3.1k).

Subject H5 used ~250ms of strong evidence to decide on each trial, but during the motor response, collected an additional ~150ms of evidence which could have informed confidence. By these measures, post-decisional evidence was potentially a far more reliable source of information with respect to choice evidence than for our other subjects. Figure A3.1L shows that post-decisional evidence *was* used, and *strongly* opposed the chosen hypothesis on low confidence trials.

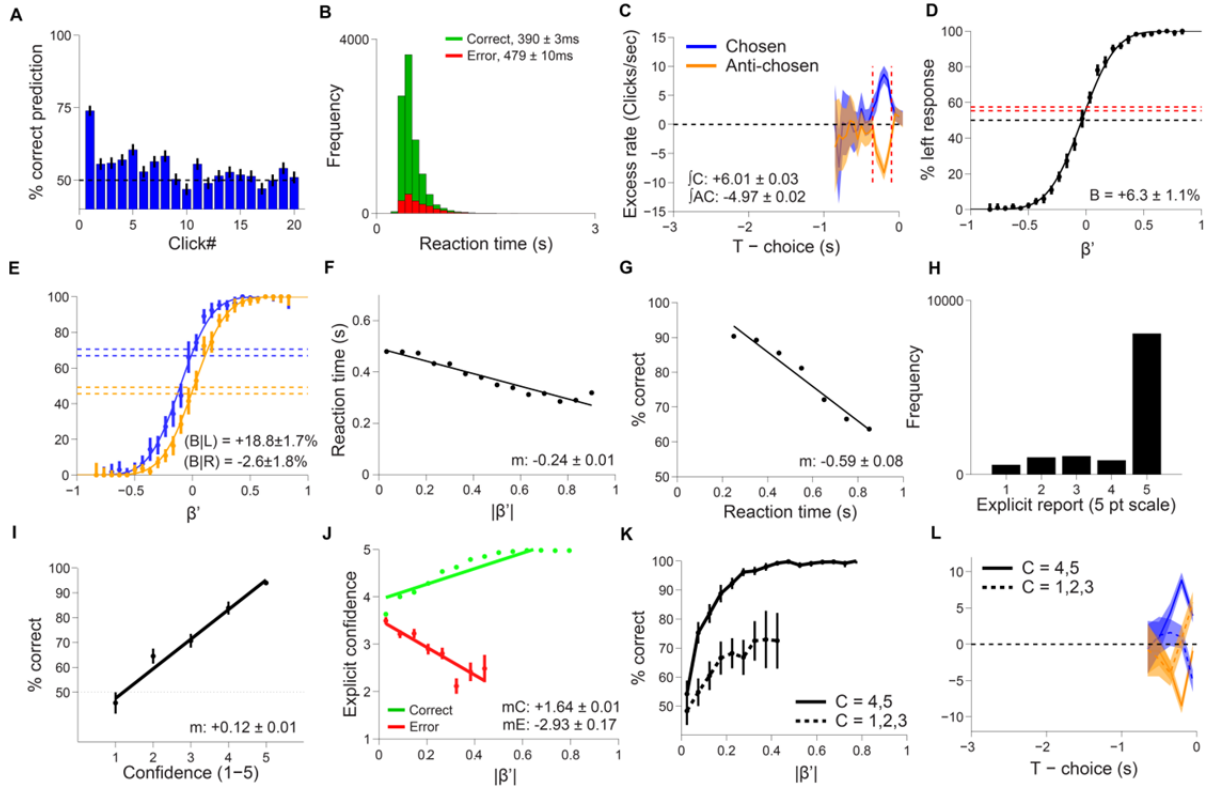


Figure A3.1: With strong evidence and biased, rapid judgments similar to our rats, normative patterns persist in explicit confidence. **A:** Subject H5 had a significant choice bias towards the side that clicked first. The seven subsequent clicks all show small but significant influence over choice. **B:** The subject responded rapidly after evidence onset but spent more time sampling on incorrect trials. Inset shows mean and standard error. **C:** Choice-aligned reverse correlation of chosen and anti-chosen evidence streams. Data were evaluated in 50ms bins. Chosen and anti-chosen streams were distinguishable by a bootstrap test for difference between means, in a window from 450ms to 150ms before choice (indicated with red dashed lines). Inset shows mean click bias calculated in the same window. Chosen-side clicks contributed to choice slightly more than anti-chosen side clicks. **D:** Psychometric function showing that evidence balance explains choice probability despite the first-click bias. A small left bias of 6% was evident. **E:** The subject showed considerable bias towards the side rewarded on the preceding trial. Biases after non-rewarded trials were less pronounced (not shown) **F:** On trials with weak evidence, the subject sampled for longer. **G:** On trials with long response times, the subject was less accurate. **H:** Subject H5 used the “5” response more often than the remaining scale divisions combined. **I:** Despite the unusual scale usage pattern, confidence reports predicted accuracy with nearly perfect calibration. **J:** Confidence on strong evidence trials was lowest when the subject erred and highest when the subject was correct. A saturation effect is evident for correct trials of very high evidence strength. **K:** The subject’s confidence predicted choice even when experimentally delivered evidence was fixed. **L:** The reverse correlation from panel C was partitioned into low and high confidence bins. A very strong bias in the post-decision period is evident on low confidence trials, for evidence that contradicted the choice.

Strikingly, in this subject, a phase of early evidence which strongly predicted choice on evidence neutral trials (450ms-300ms before choice), did not predict confidence at all (Figure A3.2). This indicates that under some circumstances, the processes of information gathering for choice and for confidence are dissociable. In chapter 5, we explore models that use post-decisional evidence exclusively, and propose a new model that can use both decision and post-decision epochs in varying proportions. In terms of our model, the fact that a period of evidence collection *only for choice* can exist in a human subject, is consistent with the idea that subjects may adapt the coupling coefficient between accumulators to the statistics of sensory evidence and reward contingencies in each task. Future research may test how flexible this trade-off is, by collecting evidence from subjects in speed and accuracy conditions, and with different distributions of evidence strength.

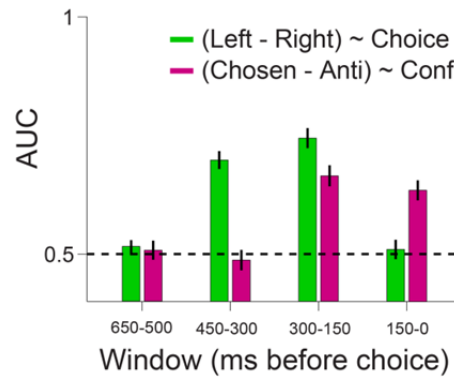


Figure A3.2: The processes of gathering evidence for choice and for confidence are dissociable. We partitioned time before choice into four bins on neutral-evidence trials. A post-decisional evidence bin was set between -150ms and 0ms, based on bootstrap tests for different means, performed on data from Figure A3.1C. The first significant bin and final significant bins for choice occurred at -150ms and -450ms. This period was divided equally into early and late choice bins. A pre-decision bin was computed between 500ms and 650ms. Clicks in subject H5's early choice epoch predicted choice but not confidence. However, evidence gathered in the final 150ms did not inform confidence *more* than evidence collected 300-150ms during the choice period, as predicted by the post-decision model's mutually exclusive trade-off between choice and confidence information on individual trials (Figure 5.5).

Appendix IV

Towards an identity test for confidence signals

It is tempting to infer, from the presence of normative model patterns in human confidence, that the human brain is computing statistical confidence from its evidence. However we cannot entirely dismiss the possibility that signals which produce the normative model patterns could exist, which are differentiable from confidence in some fundamental way. Our normative model predictions make strong predictions about the *sensitivity* of this set of patterns in identifying confidence signals. While a theoretical treatment of the specificity of these patterns is an excellent topic for future research, in this appendix I explore one instance of the ability of this set of patterns to discern between decision confidence and a superficially similar cognitive signal.

In one view, the absolute noise level corrupting perceived evidence may be reduced, or the amount of available evidence may be increased, when a channel providing evidence is *attended* to (Neill and Westberry 1987; Lu and Doshier 1998). Intuitively, these simplistic and abstract forms of “attention” should satisfy several of the normative model predictions. When attention varies on each trial, it can predict the amount of uncorrupted evidence apparent to the decision maker, and by inference, the decision maker’s accuracy. It is also intuitive that only when a decision maker is attending least, they will make mistakes in the presence of strong evidence, satisfying the error trial prediction of the second pattern. For a fixed strength of provided evidence, the strength of the same evidence after modulation should predict accuracy on average, satisfying the third prediction. However, this form of “attention”, while enhancing or suppressing the collection of information in general, does not contain direct insight into the *relative* statistics of the two samples of evidence, apparent to the decision process. We wondered whether a variable that intuitively mimics confidence in so many ways but does not have insight into the ratio of evidence supporting different hypotheses, would also satisfy the three normative predictions of decision confidence.

To test this, we constructed a two-choice decision making simulation with Gaussian evidence, identical to the one we used to evaluate two-sample t-test confidence in Chapter 3. On each of 1,000,000 trials, two samples of 30 measurements were drawn from two Gaussian distributions of variance = 1 (distributions A and B). The means of distributions A and B were offset from 0 by random amounts drawn from a uniform distribution between -0.5 and 0.5, where the means described *provided* evidence for each hypothesis (the distinction between provided and perceptual evidence discussed in section 3.1). The difference between the means of A and B on each trial described *discriminability*. I scored each trial as correct if the ordinal relationship between sample means was true of the population means based on our offsets, and incorrect if false.

We tested two attention-like “global gain” (GG) variables, which affected the strength of signals supporting separate hypotheses without directly changing their balance of evidence ratio. These two measures were: 1. *added noise*, and 2. *sample size*. These factors modulated both samples in the same way on each trial. For our added noise simulation (Figure A4.1a-d), each trial’s evidence was corrupted by adding zero-mean Gaussian noise to the evidence on each trial. The variance of the noise added (*noise variance* in Figure A4.1), ranged from 0 to 1.5 with uniform probability on each trial. T-tests were also performed on each trial’s evidence (Figure A4.1e-g). For our sample size simulation (Figure A4.1h-k), the total number of samples collected for the two hypotheses was co-modulated on a trial by trial basis by co-varying the (equivalent) sample size of the two samples being compared between 2 and 50 with equal probability.

Accuracy varied continuously with discriminability between chance and near-perfect in each simulation (Figure A4.1a,h). However, the GG measures did not span the complete range of accuracy, as did statistical confidence in our simulation (Figure A4.1b,i vs. panel e) and verbal confidence in our study. On error trials, these two GG measures were indeed lowest in the presence of strong evidence (Figure A4.1. However, GG measures were *not* positively correlated with evidence strength on correct trials as was always true for confidence reports.

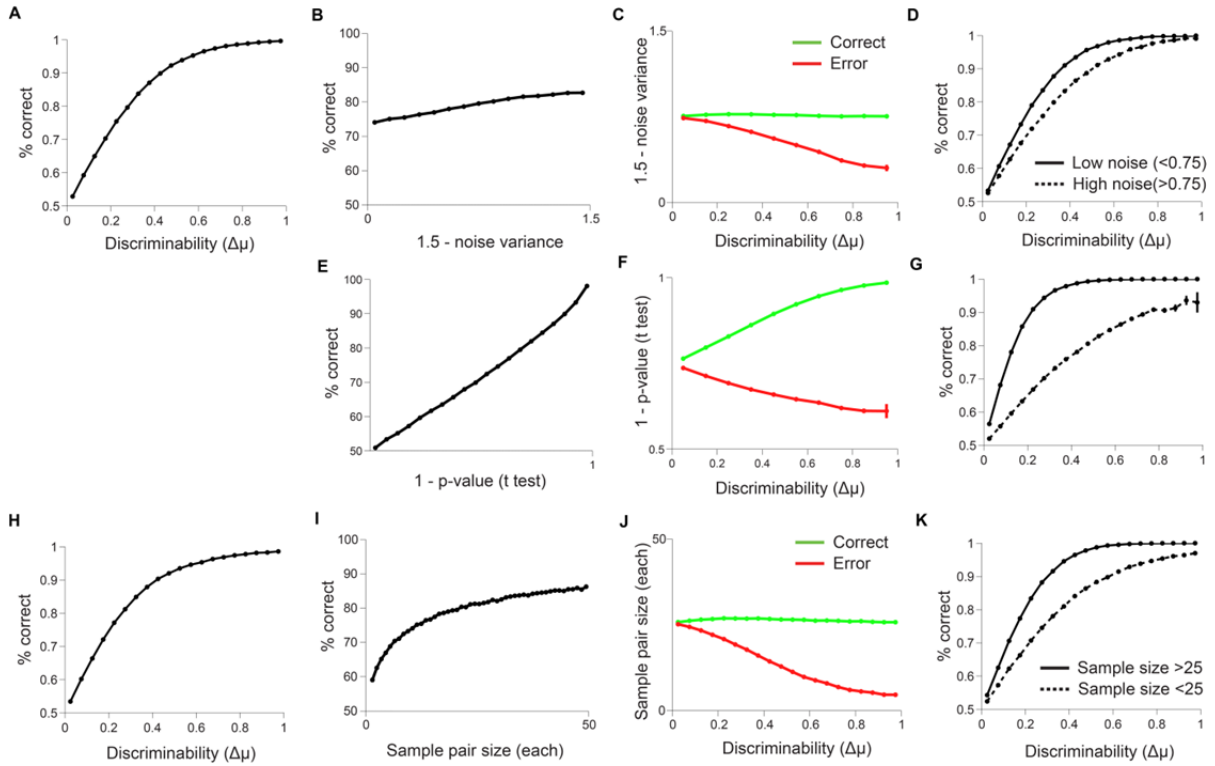


Figure A4.1: Normative predictions are not reproduced by two global gain (GG) processes. We compared two factors affecting the total amount of evidence available to a decision maker: the amount of noise added to samples of evidence (a-d) and the equivalent sample size of the two samples compared on each trial (h-k) to determine whether these modulators of evidence signal strength would be distinguishable from decision confidence using normative model patterns. **A,H:** Psychometric functions showed that accuracy increased with the difference between underlying distribution means ($\Delta\mu$), from which samples were drawn on each trial. **B,I:** Despite evidence eliciting a complete range of accuracy for the decision process, GG measures were weak accuracy predictors. **C,J:** Contrary to normative model predictions, GG measures were not positively correlated with discriminability on correct trials. **D,K:** GG measures predicted accuracy for a given difficulty. **E,F,G:** One-sided, two-sample t-tests were performed on the evidence of each trial in the *added noise* simulation, producing robust confidence patterns despite modulation of evidence.

In chapter 5, we examined two decision measures which might have been explicitly used by the brain to know its own confidence – reaction time and evidence variance – and determined that these measures could not produce normative model patterns in the context of the drift diffusion model, eluding a fit to human data. Here, we show that two instances of an additional class of decision making signal, global gain signals, do not reproduce the expected patterns of the normative confidence model, despite directly affecting the statistics of evidence used for choice. These

processes modulated evidence strength by sampling more of it, or adding noise to corrupt it, but *did not affect the balance of evidence supporting the two hypotheses on average*, preserving the primary information used to decide.

In this appendix, I showed that a set of three statistical confidence patterns can be used to discern between two related measurements of the evidence used in two-choice decision making – global gain (GG) and decision confidence. The failure of GG measures to mimic the direct correlation between provided evidence and confidence on correct trials touches on the power of this set of patterns to infer insight about the role of a candidate confidence signal in the process of decision making. The second normative pattern especially, may be an exclusive hallmark of an *a posteriori* appraisal of the likelihood that a noisy decision process chose correctly, in contrast to the pattern produced by the *a priori* global gain measure. This pattern in provided evidence requires information about the strength of evidence used to decide each trial *as it was apparent to the decision maker* (after sampling errors and all other internal noise sources were added), in order to manifest. Whether a process acquires this privileged information about a decision maker's internal evidence directly from the decision process, or tangentially as in the case of the 2DSD model (Chapter 5), a calibrated confidence signal can in principal, be used by organisms to assess the quality of their judgments and cast wagers with higher efficiency than if it were making these judgments based on a signal that modulates evidence strength prior to choice (Figure A4.1b,i). Though further research is necessary to establish whether there exists an independently generated signal that mimics the privileged insight of decision confidence in these three important ways, it may be of semantic concern whether or not a signal that exhibits the three normative patterns is in fact, a measure of decision confidence.

References

- Abramov, I., L. Hainline, et al. (1984). "Rocket-ship psychophysics. Assessing visual functioning in young children." Investigative ophthalmology & visual science **25**(11): 1307-1315.
- Adams, J. K. and P. A. Adams (1961). "Realism of confidence judgments." Psychological Review **68**(1): 33.
- Adelson, E. H. and J. R. Bergen (1985). "Spatiotemporal energy models for the perception of motion." J. Opt. Soc. Am. A **2**(2): 284-299.
- Aertsen, A. and P. Johannesma (1981). "The spectro-temporal receptive field." Biological cybernetics **42**(2): 133-143.
- Andermann, M. L., A. M. Kerlin, et al. (2010). "Chronic cellular imaging of mouse visual cortex during operant behavior and passive viewing." Frontiers in Cellular Neuroscience **4**.
- Audley, R. and A. Pike (1965). "Some alternative stochastic models of choice." British Journal of Mathematical and Statistical Psychology **18**(2): 207-225.
- Audley, R. J. (1960). "A stochastic model for individual choice behavior." Psychol Rev **67**: 1-15.
- Bar-Tal, Y., A. Sarid, et al. (2001). "A test of the overconfidence phenomenon using audio signals." The Journal of General Psychology **128**(1): 76-80.
- Baranski, J. V. and W. M. Petrusic (1994). "The calibration and resolution of confidence in perceptual judgments." Attention, Perception, & Psychophysics **55**(4): 412-428.
- Baranski, J. V. and W. M. Petrusic (1998). "Probing the locus of confidence judgments: Experiments on the time to determine confidence." Journal of Experimental Psychology: Human Perception and Performance **24**(3): 929.
- Beck, J. M., W. J. Ma, et al. (2008). "Probabilistic population codes for Bayesian decision making." Neuron **60**(6): 1142-1152.
- Behrens, T. E., M. W. Woolrich, et al. (2007). "Learning the value of information in an uncertain world." Nature neuroscience **10**(9): 1214-1221.
- Björkman, M., P. Juslin, et al. (1993). "Realism of confidence in sensory discrimination: The underconfidence phenomenon." Attention, Perception, & Psychophysics **54**(1): 75-81.
- Bodelón, C., M. Fallah, et al. (2007). "Temporal resolution for the perception of features and conjunctions." The Journal of Neuroscience **27**(4): 725-730.
- Bogacz, R., E. Brown, et al. (2006). "The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks." Psychological Review **113**(4): 700.
- Bornstein, B. H. and D. J. Zickafoose (1999). "'I know I know it, I know I saw it': The stability of the confidence-accuracy relationship across domains." Journal of Experimental Psychology: Applied **5**(1): 76.
- Brainard, D. H. (1997). "The psychophysics toolbox." Spatial vision **10**(4): 433-436.
- Bregman, A. S. (1994). Auditory scene analysis: The perceptual organization of sound, MIT press.
- Brier, G. W. (1950). "Verification of forecasts expressed in terms of probability." Monthly weather review **78**(1): 1-3.

- Brown, S. and A. Heathcote (2005). "A ballistic model of choice response time." Psychological Review **112**(1): 117-128.
- Brown, S. D. and A. Heathcote (2008). "The simplest complete model of choice response time: Linear ballistic accumulation." Cognitive psychology **57**(3): 153-178.
- Brunton, B. B. M. and Carlos D. Brody (2013). "Rats and Humans can Optimally Accumulate Evidence for Decision Making." Science(Unknown issue).
- Brunton, B. W., M. M. Botvinick, et al. (2013). "Rats and Humans Can Optimally Accumulate Evidence for Decision-Making." Science **340**(6128): 95-98.
- Busse, L., A. Ayaz, et al. (2011). "The Detection of Visual Contrast in the Behaving Mouse." The Journal of Neuroscience **31**(31): 11351-11361.
- Cattell, J. M. (1886). "The time it takes to see and name objects." Mind **11**(41): 63-65.
- Churchland, A. K., R. Kiani, et al. (2008). "Decision-making with multiple alternatives." Nat Neurosci **11**(6): 693-702.
- Clifford, C. W., E. Arabzadeh, et al. (2008). "Getting technical about awareness." Trends in cognitive sciences **12**(2): 54-58.
- Cornsweet, T. N. (1962). "The staircase-method in psychophysics." The American Journal of Psychology **75**(3): 485-491.
- Courville, A. C., N. D. Daw, et al. (2006). "Bayesian theories of conditioning in a changing world." Trends in cognitive sciences **10**(7): 294-300.
- Cury, K. M. and N. Uchida (2010). "Robust Odor Coding via Inhalation-Coupled Transient Activity in the Mammalian Olfactory Bulb." Neuron **68**(3): 570-585.
- Dawes, J. (2008). "Do Data Characteristics Change According to the number of scale points used? An experiment using 5 point, 7 point and 10 point scales." International Journal of Market Research **51**(1).
- Dombeck, D. A., A. N. Khabbaz, et al. (2007). "Imaging Large-Scale Neural Activity with Cellular Resolution in Awake, Mobile Mice." Neuron **56**(1): 43-57.
- Drugowitsch, J., R. Moreno-Bote, et al. (2012). "The cost of accumulating evidence in perceptual decision making." The Journal of Neuroscience **32**(11): 3612-3628.
- Eagleman, D. M., U. T. Peter, et al. (2005). "Time and the brain: how subjective time relates to neural time." The Journal of Neuroscience **25**(45): 10369-10371.
- Efron, B. and R. Tibshirani (1993). An introduction to the bootstrap, CRC press.
- Erlich, Jeffrey C., M. Bialek, et al. (2011). "A Cortical Substrate for Memory-Guided Orienting in the Rat." Neuron **72**(2): 330-343.
- Felsen, G. and Z. F. Mainen (2008). "Neural Substrates of Sensory-Guided Locomotor Decisions in the Rat Superior Colliculus." Neuron **60**(1): 137-148.
- Ferrell, W. R. and P. J. McGoey (1980). "A model of calibration for subjective probabilities." Organizational Behavior and Human Performance **26**(1): 32-53.
- Fimbel, E. J., R. Michaud, et al. (2009). "Certainty in categorical judgment of size." PloS one **4**(7): e6198.
- Fine, I. (2009). "Fitting Psychometric Functions." Introduction to Programming for the Behavioral Sciences Retrieved 8/3/2013, 2013, from http://courses.washington.edu/matlab1/Lesson_5.html.
- Flavell, J. H. (1979). "Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry." American psychologist **34**(10): 906.

- Fleming, S. M. and R. J. Dolan (2010). "Effects of loss aversion on post-decision wagering: Implications for measures of awareness." Consciousness and cognition **19**(1): 352-363.
- Fleming, S. M., R. S. Weil, et al. (2010). "Relating introspective accuracy to individual differences in brain structure." Science **329**(5998): 1541-1543.
- Foote, A. L. and J. D. Crystal (2007). "Metacognition in the Rat." Current Biology **17**(6): 551-555.
- Gold, J. I. and M. N. Shadlen (2007). "The Neural Basis of Decision Making." Annual Review of Neuroscience **30**(1): 535-574.
- Green, D. M. and J. A. Swets (1966). SIGNAL DETECTION THEORY AND PSYCHOPHYSICS. Oxford, England, John Wiley.
- Griffin, D. and A. Tversky (1992). "The weighing of evidence and the determinants of confidence." Cognitive psychology **24**(3): 411-435.
- Hampton, R. R. (2001). "Rhesus monkeys know when they remember." Proceedings of the National Academy of Sciences **98**(9): 5359-5362.
- Harvey, C. D., P. Coen, et al. (2012). "Choice-specific sequences in parietal cortex during a virtual-navigation decision task." Nature **484**(7392): 62-68.
- Harvey, C. D., F. Collman, et al. (2009). "Intracellular dynamics of hippocampal place cells during virtual navigation." Nature **461**(7266): 941-946.
- Heathcote, A., S. Brown, et al. (2002). "Quantile maximum likelihood estimation of response time distributions." Psychonomic bulletin & review **9**(2): 394-401.
- Hebb, D. (1949). "The organization of behavior; a neuropsychological theory."
- Henmon, V. A. C. (1911). "The relation of the time of a judgment to its accuracy." Psychological Review **18**(3): 186.
- Histed, M. H., L. A. Carvalho, et al. (2012). "Psychophysical measurement of contrast sensitivity in the behaving mouse." Journal of neurophysiology **107**(3): 758-765.
- Holt, G. R., W. R. Softky, et al. (1996). "Comparison of discharge variability in vitro and in vivo in cat visual cortex neurons." Journal of neurophysiology **75**(5): 1806-1814.
- Inman, A. and S. J. Shettleworth (1999). "Detecting metamemory in nonverbal subjects: A test with pigeons." Journal of Experimental Psychology: Animal Behavior Processes **25**(3): 389.
- Insabato, A., M. Pannunzi, et al. (2010). "Confidence-related decision making." Journal of neurophysiology **104**(1): 539-547.
- Jastrow, J. and C. Peirce (1884). "On small differences in sensation." Memoirs of the National Academy of Science **3**: 1884.
- Jonsson, A.-C. and C. M. Allwood (2003). "Stability and variability in the realism of confidence judgments over time, content domain, and gender." Personality and Individual Differences **34**(4): 559-574.
- Jun, J., B. W. Brunton, et al. (2010). Automated tracking of rat body position and orientation during performance of decision making tasks in high-throughput training systems. Program No. 805.14/JJJ40 Neuroscience Meeting Planner. Society for Neuroscience 2010. San Diego, CA.
- Kahneman, D. and A. Tversky (1972). "Subjective probability: A judgment of representativeness." Cognitive psychology **3**(3): 430-454.
- Kahneman, D. and A. Tversky (1979). "Prospect theory: An analysis of decision under risk." Econometrica: Journal of the Econometric Society: 263-291.

- Kepecs, A. and Z. F. Mainen (2012). "A computational framework for the study of confidence in humans and animals." Philosophical Transactions of the Royal Society B: Biological Sciences **367**(1594): 1322-1337.
- Kepecs, A., N. Uchida, et al. (2008). "Neural correlates, computation and behavioural impact of decision confidence." Nature **455**(7210): 227-231.
- Keren, G. (1988). "On the ability of monitoring non-veridical perceptions and uncertain knowledge: Some calibration studies." Acta Psychologica **67**(2): 95-119.
- Keren, G. (1991). "Calibration and probability judgements: Conceptual and methodological issues." Acta Psychologica **77**(3): 217-273.
- Kiani, R., T. D. Hanks, et al. (2008). "Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment." The Journal of Neuroscience **28**(12): 3017-3029.
- Kiani, R. and M. N. Shadlen (2009). "Representation of confidence associated with a decision by neurons in the parietal cortex." Science **324**(5928): 759-764.
- Knill, D. C. and A. Pouget (2004). "The Bayesian brain: the role of uncertainty in neural coding and computation." Trends in Neurosciences **27**(12): 712-719.
- Koehler, D. J. and N. Harvey (1997). "Confidence judgments by actors and observers." Journal of Behavioral Decision Making **10**(3): 221-242.
- Komiyama, T., T. R. Sato, et al. (2010). "Learning-related fine-scale specificity imaged in motor cortex circuits of behaving mice." Nature **464**(7292): 1182-1186.
- Kvavilashvili, S. and W. Koutstaal (2008). "Confidence and decision type under matched stimulus conditions: overconfidence in perceptual but not conceptual decisions." Journal of Behavioral Decision Making **21**(3): 253-281.
- LaBerge, D. (1962). "A recruitment theory of simple behavior." Psychometrika **27**(4): 375-396.
- Lagarias, J. C., J. A. Reeds, et al. (1998). "Convergence properties of the Nelder--Mead simplex method in low dimensions." SIAM Journal on Optimization **9**(1): 112-147.
- Li, Z., J. Burrone, et al. (2005). "Synaptic vesicle recycling studied in transgenic mice expressing synaptobrevin." Proceedings of the National Academy of Sciences **102**(17): 6131-6136.
- Lichtenstein, S. and B. Fischhoff (1977). "Do those who know more also know more about how much they know?" Organizational Behavior and Human Performance **20**(2): 159-183.
- Lichtenstein, S. and B. Fischhoff (1980). "Training for calibration." Organizational Behavior and Human Performance **26**(2): 149-171.
- Lichtenstein, S., B. Fischhoff, et al. (1981). Calibration of probabilities: The state of the art to 1980, DTIC Document.
- Likert, R. (1932). "A technique for the measurement of attitudes." Archives of psychology.
- Lindeberg, J., D. Usoskin, et al. (2004). "Transgenic expression of Cre recombinase from the tyrosine hydroxylase locus." genesis **40**(2): 67-73.
- Link, S. W. and R. A. Heath (1975). "A sequential theory of psychological discrimination." Psychometrika **40**(1): 77-105.
- Lu, Z.-L. and B. A. Doshier (1998). "External noise distinguishes attention mechanisms." Vision Research **38**(9): 1183-1198.
- MacLeod, C. M. and T. O. Nelson (1984). "Response latency and response accuracy as measures of memory." Acta Psychologica **57**(3): 215-235.

- Manly, B. F. (2007). Randomization, bootstrap and Monte Carlo methods in biology, CRC Press.
- Mazurek, M. E., J. D. Roitman, et al. (2003). "A Role for Neural Integrators in Perceptual Decision Making." Cerebral Cortex **13**(11): 1257-1269.
- McClelland, J. L. (2001). "The time course of perceptual choice: The leaky, competing accumulator model." Psychological Review **108**(3): 550-592.
- Mehta, S. B., D. Whitmer, et al. (2007). Active Spatial Perception in the Vibrissa Scanning Sensorimotor System, Public Library of Science.
- Merkle, E. C. and T. Van Zandt (2006). "An application of the Poisson race model to confidence calibration." Journal of Experimental Psychology: General **135**(3): 391.
- Metcalfe, J. E. and A. P. Shimamura (1994). Metacognition: Knowing about knowing, The MIT Press.
- Middlebrooks, P. G. and M. A. Sommer (2011). "Metacognition in monkeys during an oculomotor task." Journal of Experimental Psychology-Learning Memory and Cognition **37**(2): 325.
- Milosavljevic, M., J. Malmaud, et al. (2010). "The drift diffusion model can account for value-based choice response times under high and low time pressure." Judgement & Decision Making **5**: 437-449.
- Moreno-Bote, R. (2010). "Decision confidence and uncertainty in diffusion models with partially correlated neuronal integrators." Neural computation **22**(7): 1786-1811.
- Neill, W. T. and R. L. Westberry (1987). "Selective attention and the suppression of cognitive noise." Journal of Experimental Psychology: Learning, Memory, and Cognition **13**(2): 327.
- Neri, P., A. J. Parker, et al. (1999). "Probing the human stereoscopic system with reverse correlation." Nature **401**(6754): 695-698.
- Newsome, W. T., K. H. Britten, et al. (1989). "Neuronal correlates of a perceptual decision." Nature **341**(6237): 52-54.
- Nienborg, H. and B. G. Cumming (2009). "Decision-related activity in sensory neurons reflects more than a neuron's causal effect." Nature **459**(7243): 89-92.
- Olsson, H. and A. Winman (1996). "Underconfidence in sensory discrimination: The interaction between experimental setting and response strategies." Attention, Perception, & Psychophysics **58**(3): 374-382.
- Otazu, G. H., L.-H. Tai, et al. (2009). "Engaging in an auditory task suppresses responses in auditory cortex." Nat Neurosci **12**(5): 646-654.
- Palmer, J., A. C. Huk, et al. (2005). "The effect of stimulus strength on the speed and accuracy of a perceptual decision." J Vis **5**(5): 376-404.
- Peirce, C. S. (1877). "Illustrations of the logic of science. Fourth paper - the probability of induction. ." The Popular Science Monthly **12**: 705-718.
- Peirce, J. W. (2007). "PsychoPy—psychophysics software in Python." Journal of Neuroscience Methods **162**(1): 8-13.
- Persaud, N., P. McLeod, et al. (2007). "Post-decision wagering objectively measures awareness." Nat Neurosci **10**(2): 257-261.
- Peterson, W., T. Birdsall, et al. (1954). "The theory of signal detectability." Information Theory, IRE Professional Group on **4**(4): 171-212.

- Petrusic, W. M. and J. V. Baranski (2003). "Judging confidence influences decision processing in comparative judgments." Psychonomic bulletin & review **10**(1): 177-183.
- Pleskac, T. J. and J. R. Busemeyer (2010). "Two-stage dynamic signal detection: a theory of choice, decision time, and confidence." Psychological Review **117**(3): 864.
- Rahnev, D. A., B. Maniscalco, et al. (2012). "Direct injection of noise to the visual cortex decreases accuracy but increases decision confidence." Journal of neurophysiology **107**(6): 1556-1563.
- Ramamritham, K., C. Shen, et al. (1998). Using Windows NT for real-time applications: Experimental observations and recommendations. Real-Time Technology and Applications Symposium, 1998. Proceedings. Fourth IEEE, IEEE.
- Rao, R. P. N. (2010). "Decision Making under Uncertainty: A Neural Model based on Partially Observable Markov Decision Processes." Frontiers in Computational Neuroscience **4**.
- Ratcliff, R. (1978). "A theory of memory retrieval." Psychological Review **85**(2): 59-108.
- Ratcliff, R. (1988). "Continuous versus discrete information processing: Modeling accumulation of partial information." Psychological Review **95**(2): 238-255.
- Ratcliff, R., Y. T. Hasegawa, et al. (2007). "Dual Diffusion Model for Single-Cell Recording Data From the Superior Colliculus in a Brightness-Discrimination Task." Journal of neurophysiology **97**(2): 1756-1774.
- Ratcliff, R. and G. McKoon (2008). "The diffusion decision model: Theory and data for two-choice decision tasks." Neural computation **20**(4): 873-922.
- Ratcliff, R. and J. N. Rouder (1998). "Modeling response times for two-choice decisions." Psychological Science **9**(5): 347-356.
- Ratcliff, R. and P. L. Smith (2004). "A comparison of sequential sampling models for two-choice reaction time." Psychological Review **111**(2): 333.
- Resulaj, A., R. Kiani, et al. (2009). "Changes of mind in decision-making." Nature **461**(7261): 263-266.
- Rinberg, D., A. Koulakov, et al. (2006). "Speed-Accuracy Tradeoff in Olfaction." Neuron **51**(3): 351-358.
- Ringach, D. and R. Shapley (2004). "Reverse correlation in neurophysiology." Cognitive Science **28**(2): 147-166.
- Rockette, H. E., D. Gur, et al. (1992). "The use of continuous and discrete confidence judgments in receiver operating characteristic studies of diagnostic imaging techniques." Investigative radiology **27**(2): 169-172.
- Roitman, J. D. and M. N. Shadlen (2002). "Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task." J Neurosci **22**(21): 9475-9489.
- Sanders, J. and A. Kepecs (2012). "Choice Ball: a response interface for psychometric discrimination in head-fixed mice." Journal of neurophysiology.
- Schall, J. D. and K. G. Thompson (1999). "Neural selection and control of visually guided eye movements." Annu Rev Neurosci **22**: 241-259.
- Schapire, R. E. and Y. Singer (1999). "Improved boosting algorithms using confidence-rated predictions." Machine learning **37**(3): 297-336.
- Schwarz, C., H. Hentschke, et al. (2010). "The head-fixed behaving rat—Procedures and pitfalls." Somatosensory & Motor Research **27**(4): 131-148.

- Schwenk, H. and Y. Bengio (2000). "Boosting neural networks." Neural computation **12**(8): 1869-1887.
- Shields, W. E., J. D. Smith, et al. (1997). "Uncertain responses by humans and Rhesus monkeys (*Macaca mulatta*) in a psychophysical same-different task." Journal of Experimental Psychology: General **126**(2): 147.
- Simen, P., D. Contreras, et al. (2009). "Reward rate optimization in two-alternative decision making: empirical tests of theoretical predictions." Journal of experimental psychology. Human perception and performance **35**(6): 1865.
- Skinner, B. F. (1938). "The behavior of organisms: An experimental analysis."
- Smith, J. D., M. J. Beran, et al. (2006). "Dissociating uncertainty responses and reinforcement signals in the comparative study of uncertainty monitoring." Journal of Experimental Psychology-General **135**(2): 282-297.
- Smith, J. D., J. Schull, et al. (1995). "The uncertain response in the bottlenosed dolphin (*Tursiops truncatus*)." Journal of Experimental Psychology: General **124**(4): 391.
- Smith, J. D., W. E. Shields, et al. (1997). "The uncertain response in humans and animals." Cognition **62**(1): 75-97.
- Smith, P. L. and D. Vickers (1988). "The accumulator model of two-choice discrimination." Journal of Mathematical Psychology **32**(2): 135-168.
- Soderquist, D. R. and R. D. Shilling (1992). "Forward masking in young children: Rocketship psychophysics." Current Psychology **11**(2): 145-156.
- Sole, L. M., S. J. Shettleworth, et al. (2003). "Uncertainty in pigeons." Psychonomic bulletin & review **10**(3): 738-745.
- Sollich, P. (2002). "Bayesian methods for support vector machines: Evidence and predictive class probabilities." Machine learning **46**(1): 21-52.
- Stanford, T. R., S. Shankar, et al. (2010). "Perceptual decision making in less than 30 milliseconds." Nature neuroscience **13**(3): 379-385.
- Stankov, L. (1998). "Calibration curves, scatterplots and the distinction between general knowledge and perceptual tasks." Learning and Individual Differences **10**(1): 29-50.
- Stuttgen, M. C. and C. Schwarz (2008). "Psychophysical and neurometric detection performance under stimulus uncertainty." Nat Neurosci **11**(9): 1091-1099.
- Sun, R., M. B. Bouchard, et al. (2010). A Low-Cost, Portable System for High-Speed Multispectral Optical Imaging. Biomedical Optics, Optical Society of America.
- Sutton, J. E. and S. J. Shettleworth (2008). "Memory without awareness: pigeons do not show metamemory in delayed matching to sample." Journal of Experimental Psychology: Animal Behavior Processes **34**(2): 266.
- Sutton, R. S. and A. G. Barto (1998). Reinforcement learning: An introduction, Cambridge Univ Press.
- Tanahira, C., S. Higo, et al. (2009). "Parvalbumin neurons in the forebrain as revealed by parvalbumin-Cre transgenic mice." Neuroscience Research **63**(3): 213-223.
- Taniguchi, H., M. He, et al. (2011). "A Resource of Cre Driver Lines for Genetic Targeting of GABAergic Neurons in Cerebral Cortex." Neuron **71**(6): 995-1013.
- Teikari, P., R. P. Najjar, et al. (2012). "An inexpensive Arduino-based LED stimulator system for vision research." Journal of Neuroscience Methods.
- Terrace, H. S. (2004). The Missing Link in Cognition: Origins of Self-Reflective Consciousness: Origins of Self-Reflective Consciousness, Oxford University Press.

- Usher, M. and J. L. McClelland (2001). "The time course of perceptual choice: the leaky, competing accumulator model." Psychological Review **108**(3): 550.
- Usher, M. and D. Zakay (1993). "A neural network model for attribute-based decision processes." Cognitive Science **17**(3): 349-396.
- Van Zandt, T., H. Colonius, et al. (2000). "A comparison of two response time models applied to perceptual matching." Psychonomic bulletin & review **7**(2): 208-256.
- Vickers, D. (1979). Decision processes in visual perception. New York, London, Academic Press.
- Vickers, D. (2001). Where does the balance of evidence lie with respect to confidence. Proceedings of the seventeenth annual meeting of the international society for psychophysics.
- Vickers, D. and J. Packer (1982). "Effects of alternating set for speed or accuracy on response time, accuracy and confidence in a unidimensional discrimination task." Acta Psychologica **50**(2): 179-197.
- Volkman, J. (1934). "The relation of the time of judgment to the certainty of judgment." Psychological Bulletin **31**: 672-673.
- Wald, A. (1947). Sequential analysis, Courier Dover Publications.
- Watson, C. S., S. C. Kellogg, et al. (1973). "The uncertain response in detection-oriented psychophysics." Journal of Experimental Psychology **99**(2): 180.
- Wichmann, F. A. and N. J. Hill (2001). "The psychometric function: I. Fitting, sampling, and goodness of fit." Perception & psychophysics **63**(8): 1293-1313.
- Wilimzig, C., N. Tsuchiya, et al. (2008). "Spatial attention increases performance but not subjective confidence in a discrimination task." J Vis **8**(5).
- Winman, A., P. Juslin, et al. (1998). "The confidence-hindsight mirror effect in judgment: An accuracy-assessment model for the knew-it-all-along phenomenon." Journal of Experimental Psychology: Learning, Memory, and Cognition **24**(2): 415.
- Yeung, N. and C. Summerfield (2012). "Metacognition in human decision-making: confidence and error monitoring." Philosophical Transactions of the Royal Society B: Biological Sciences **367**(1594): 1310-1321.
- Yoshida, T. and D. B. Katz (2011). "Control of Prestimulus Activity Related to Improved Sensory Coding within a Discrimination Task." The Journal of Neuroscience **31**(11): 4101-4112.
- Zariwala, H. A., B. G. Borghuis, et al. (2012). "A Cre-dependent GCaMP3 reporter mouse for neuronal imaging in vivo." J Neurosci **32**(9): 3131-3141.
- Znamenskiy, P. and A. M. Zador (2013). "Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination." Nature **497**(7450): 482-485.
- Zylberberg, A., P. Barttfeld, et al. (2012). "The construction of confidence in a perceptual decision." Frontiers in Integrative Neuroscience **6**.